UNIVERSIDAD DE CÓRDOBA

Programa de doctorado

Ingeniería Agraria, Alimentaria, Forestal y de Desarrollo Rural Sostenible.

TESIS DOCTORAL

Identification and Development of Cereals Suitable for People suffering from Gluten Allergies and Intolerances: Directed Mutagenesis by Specific Nucleases (CRISPR/Cas9) of Immunodominant Genes in Relation to Celiac Disease.

Identificación y Desarrollo de Cereales Aptos para Personas con Alergias e Intolerancia al Gluten: Mutagénesis Dirigida mediante Nucleasas Específicas (CRISPR/Cas9) de Genes Inmunodominantes en Relación a la Celiaquía.

Autora
Susana Sánchez León

Director
Dr. Francisco Barro Losada

Octubre 2019

Tesis Financiada por los proyectos AGL2013-48946-C3-1-R y AGL2016-80566-P del Ministerio de Economía y Competitividad y por el Fondo Europeo de Desarrollo Regional (FEDER) y realizada en el Instituto de Agricultura Sostenible-CSIC

CSIC Consejo Superior de Investigaciones Científicas | Instituto de Agricultura Sostenible IAS

TITULO: *Identification and Development of Cereals Suitable for People suffering from Gluten Allergies and Intolerances: Directed Mutagenesis by Specific Nucleases (CRISPR/Cas9) of Immunodominant Genes in Relation to Celiac Disease*

AUTOR: *Susana Sánchez León*

© Edita: UCOPress. 2020
Campus de Rabanales
Ctra. Nacional IV, Km. 396 A
14071 Córdoba

https://www.uco.es/ucopress/index.php/es/
ucopress@uco.es

TÍTULO DE LA TESIS: Identification and Development of Cereals Suitable for People suffering from Gluten Allergies and Intolerances: Directed Mutagenesis by Specific Nucleases (CRISPR/Cas9) of Immunodominant Genes in Relation to Celiac Disease.

DOCTORANDO/A: Susana Sánchez León

INFORME RAZONADO DEL/DE LOS DIRECTOR/ES DE LA TESIS

(Se hará mención a la evolución y desarrollo de la tesis, así como a trabajos y publicaciones derivados de la misma).

El Dr. **Francisco Barro Losada**, Profesor de Investigación del Instituto de Agricultura Sostenible (IAS), perteneciente al CSIC, como director de la tesis doctoral con título: "Identification and Development of Cereals Suitable for People suffering from Gluten Allergies and Intolerances: Directed Mutagenesis by Specific Nucleases (CRISPR-Cas9) of Immunodominant Genes in Relation to Celiac Disease", realizada por **Susana Sánchez León**,

INFORMA

Que dicha tesis ha sido realizada bajo mi dirección. Su principal objetivo ha sido la identificación y desarrollo de cereales aptos para personas con alergias e intolerancia al gluten, además de la aplicación de las más modernas técnicas de mutagénesis dirigida mediante nucleasas específicas (tecnología CRISPR/Cas9) en genes inmunodominantes de la enfermedad celíaca y la caracterización de las líneas generadas.

Los resultados y conclusiones obtenidos son de gran relevancia para el desarrollo de cereales aptos para intolerancias al gluten, representa un gran avance en el conocimiento científico en cuanto a la inmunogénesis de los cereales y diversas proteínas del gluten de trigo, y supone un avance sin precedentes en cuanto a la aplicación de técnicas de mutagénesis dirigida en trigo mediante la edición genética de múltiples genes inmunodominantes.

En su etapa pre-doctoral, Susana Sánchez León realizó una estancia de 4 meses desde el 1 de Julio hasta el 1 de Noviembre de 2016, en el departamento de "Genetics, Cell Biology, and Development" perteneciente a la Universidad de Minnesota, EEUU, bajo la tutela del Profesor Daniel F. Voytas. El objetivo principal de la estancia fue adquirir nuevas habilidades y conocimientos en la aplicación de técnicas de mutagénesis dirigida mediante nucleasas específicas y de este modo impulsar el desarrollo del proyecto de investigación en el que se enmarca la tesis doctoral. Durante el trascurso de la estancia breve en la Universidad de Minnesota se diseñaron y construyeron un buen número de vectores de expresión CRISPR/Cas9 para la ingeniería genética de plantas.

La doctoranda, además, a lo largo del desarrollo de la tesis ha colaborado en varias líneas de investigación y asistido a diversos congresos nacionales e internacionales de los que derivan las publicaciones siguientes:

Publicaciones:

Javier Gil-Humanes, Yanpeng Wang, Zhen Liang, Qiwei Shan, Carmen V. Ozuna, **Susana Sánchez-León**, Nicholas J. Baltes, Colby Starker, Francisco Barro, Caixia Gao, Daniel F. Voytas. (2016) High efficiency gene targeting in hexaploid wheat using DNA replicons and CRISPR/Cas9. **The Plant Journal**, 89: 1251-1262. Doi:10.1111/tpj.13446
Impact Factor: 5.901
ISI Journal Citation Reports Ranking 2016: 13/212 (Plant Sciences)

Susana Sánchez-León, Javier Gil-Humanes, Carmen V. Ozuna, María J. Giménez, Carolina Sousa, Daniel F. Voytas, Francisco Barro. (2017) Low-gluten, non-transgenic wheat engineered with CRISPR/Cas9. **Plant Biotechnology Journal**, 16: 902-910. Doi:10.1111/pbi.12837
Impact Factor: 6.305
ISI Journal Citation Reports Ranking 2017: 9/223 (Plant Sciences)

María Dolores García-Molina, María José Giménez, **Susana Sánchez-León**, Francisco Barro. (2019) Gluten Free Wheat: Are We There? **Nutrients**, 11: 487. Doi: 10.3390/nu11030487
Impact Factor (2018): 4.171
ISI Journal Citation Reports Ranking 2018: 16/87 (Nutrition and dietetics)

Carmen Haro, Myriam Villatoro, Luis Vaquero, Jorge Pastor, María J. Giménez, Carmen V. Ozuna, **Susana Sánchez-León**, María D. García-Molina, Verónica Segura, Isabel Comino, Carolina Sousa, Santiago Vivas, Blanca B. Landa and Francisco Barro. (2018) The Dietary Intervention of Transgenic Low-Gliadin Wheat Bread in Patients with Non-Celiac Gluten Sensitivity (NCGS) Showed No Differences with Gluten Free Diet (GFD) but Provides Better Gut Microbiota Profile. **Nutrients**, 10, 1-15. Doi:10.1111/pbi.12837
Impact Factor (2018): 4.171
ISI Journal Citation Reports Ranking 2018: 16/87 (Nutrition and dietetics)

Susana Sánchez-León, Carmen Haro, Myriam Villatoro, Luis Vaquero, Isabel Comino, Ana B. González-Amigo, Santiago Vivas, Jorge Pastor, Carolina Sousa, Blanca B. Landa and Francisco Barro. (2020)Tritordeum Breads are well Tolerated with Preference over Gluten-Free Breads in Non-Celiac Gluten-Sensitive (NCGS) Patients and its Consumption Induce Changes in Beneficial Butyrate-Producing Gut Bacteria. En revision en **Nutrients**.

Susana Sánchez-León, María José Giménez, Isabel Comino, Carolina Sousa, Miguel Ángel López Casado, María Isabel Torres, and Francisco Barro. (2020) Stimulatory Response of Celiac Disease Peripheral Blood Mononuclear Cells (PBMCs) induced by RNAi wheat lines differing in the grain proteins composition. En revision en **Nutrients.**

Aportaciones a congresos:

Susana Sánchez-León, Javier Gil-Humanes, Carmen V. Ozuna, María J. Giménez, Carolina Sousa, Daniel F. Voytas, Francisco Barro. Low-gluten wheat engineered with CRISPR/Cas9. Póster científico. Plant and Molecular Biology. Paris, France, 22-24 de Febrero de 2018.

Susana Sánchez-León, Javier Gil-Humanes, Carmen V. Ozuna, María J. Giménez, Carolina Sousa, Daniel F. Voytas, Francisco Barro. Engineering the major coeliac disease immunogenic complex in wheat by CRISPR/Cas9. Póster científico Simposio Español de Fisiología y Mejora de Cereales (SEFIMEC). Zaragoza, España, 9-10 de Abril de 2018.

Susana Sánchez-León, Javier Gil-Humanes, Carmen V. Ozuna, María J. Giménez, Carolina Sousa, Daniel F. Voytas, Francisco Barro. Engineering the coeliac disease genes in Durum and Bread wheat by CRISPR/Cas9. Comunicación Oral. From Seed to Pasta III. Bolonia, Italia, 19-21 de Septiembre de 2018.

Susana Sánchez-León, Javier Gil-Humanes, Carmen V. Ozuna, María J. Giménez, Carolina Sousa, Daniel F. Voytas, Francisco Barro. Edición genética del principal complejo inmunogénicos de la enfermedad celiaca en trigo mediante CRISPR/Cas9. Comunicación Oral. VI Congreso nacional de la Sociedad Española de Enfermedad Celíaca (SEEC). Córdoba, España, 28-30 de Noviembre de 2018.

Susana Sánchez-León, Carmen V. Ozuna, María D. García-Molina, María J. Giménez, Francisco Barro. Biotecnología y trigo (y otros cereales) sin gluten: ¿dónde estamos? Gastroenterología y hepatología. (Ed. impr.) 2018, 23-24.

Susana Sánchez-León, María D. García-Molina, María J. Giménez, Francisco Barro. Effect of nitrogen fertilization on the major α-gliadin immunogenic complex in two low-gliadin wheat lines. Póster científico y comunicación oral “flash”. II Simposio Español de Fisiología y Mejora de Cereales (SEFIMEC). Córdoba, España, 6 y 7 de Marzo de 2019.

Actividades formativas del programa de doctorado:

Participación en las actividades de mejora de la empleabilidad organizadas por la Universidad de Córdoba: Taller para la mejora de la empleabilidad dirigido por APROINN CONSULTORES. Realizadas en Córdoba, el 27 de Noviembre de 2015.

Asistencia y exposición de seminario titulado “Variantes alélicas de la celiaquía en trigo: Mutagénesis dirigida mediante nucleasas específicas de los genes inmunodominantes” en el VI Encuentro de Estudiantes de Doctorado del IAS-CSIC, celebrado en Córdoba el pasado 24 de Junio de 2016.

Asistencia al “V congreso nacional de la Sociedad Española de Enfermedad Celíaca”, (SEEC), celebrado en Barcelona del 17 al 19 de Noviembre de 2016.

Asistencia al seminario de actualidad titulado “Figuras de Protección del Medio Ambiente”, celebrado en la Universidad de Córdoba el pasado 9 de marzo de 2018.

Asistencia a la jornada técnica agroalimentaria titulada “Innovación como Motor de Negocio en Producción Vegetal”, organizado por el IAS, IFAPA y la empresa BIOVEGEN y celebrada en el IFAPA, Córdoba, el pasado 13 de marzo de 2018

Asistencia al seminario de actualidad titulado “Embalajes para productos hortícolas en fresco: abordaje térmico y estructural”, celebrado en la Universidad de Córdoba el pasado 27 de marzo de 2018.

Visitas a centros, grandes instalaciones y laboratorios de investigación del sector privado y del público: Visita a la Estación Experimental del Zaidín (EEZ). Organiza por la Universidad de Córdoba, en Granada el pasado 29 de Marzo de 2019.

Organización del Seminario de Actualidad “La Tesis Doctoral y las Oportunidades en Materia de Transferencia Tecnológica” realizado en el campus de Rabanales, Universidad de Córdoba, el pasado 16 de mayo de 2019.

Por todo ello, se autoriza la presentación de la tesis doctoral.

Córdoba, 14 de Octubre de 2019

BARRO LOSADA FRANCISCO - 30474943N

Firmado digitalmente por BARRO LOSADA FRANCISCO - 30474943N
Fecha: 2019.10.14 11:33:59 +02'00'

Firma del director

Fdo.: Francisco Barro Losada

TÍTULO DE LA TESIS: Identification and Development of Cereals Suitable for People suffering from Gluten Allergies and Intolerances: Directed Mutagenesis by Specific Nucleases (CRISPR/Cas9) of Immunodominant Genes in Relation to Celiac Disease.

DOCTORANDO/A: Susana Sánchez León

INFORME RAZONADO DEL TUTOR

(Ratificando el informe favorable del director. Sólo cuando el director no pertenezca a la Universidad de Córdoba).

Como he señalado en informes anteriores y una vez visto la memoria de la Tesis Doctoral y el informe de su director, entiendo que la Doctoranda ha cubierto objetivos propuestos, tanto desde un punto de vista académico como investigador. Ello queda reflejado en los trabajos publicados y presentaciones a Congresos. Señalaría la labor realizada y la formación adquirida como excelente.

Por todo ello, se autoriza la presentación de la tesis doctoral.

Córdoba, 15____ de octubre_____ de 2019_____

Firma del responsable de línea de investigación

JORRIN NOVO JESUS VALENTIN - 30446605X

Firmado digitalmente por JORRIN NOVO JESUS VALENTIN - 30446605X
Nombre de reconocimiento (DN): c=ES, serialNumber=IDCES-30446605X, givenName=JESUS VALENTIN, sn=JORRIN NOVO, cn=JORRIN NOVO JESUS VALENTIN - 30446605X
Fecha: 2019.10.15 10:53:29 +02'00'

Fdo.: _Prof. Jesús V. Jorrín Novo_____

Contents

Acknowledgments

El desarrollo de todo el trabajo que culmina en esta tesis doctoral es gracias al esfuerzo y compromiso de un gran grupo de personas que de una manera u otra han contribuido, a las que después de tantos años me gustaría expresar mi más sincera gratitud y enhorabuena por toda la labor realizada. Del mismo modo, me gustaría mencionar mi gratitud a los programas de "Ayudas para contratos predoctorales para la formación de doctores" (BES-2014-069009) del Ministerio de Economía, Industria y Competitividad, gracias a las cuales pude entrar a formar parte del desarrollo de los proyectos AGL2013-48946-C3-1-R y AGL2016-80566-P y del grupo de trabajo con el que he tenido la suerte de compartir estos últimos 5 años.

Especialmente me gustaría expresar mi enorme agradecimiento a mi director de tesis, Francisco Barro, la principal persona detrás de este trabajo. No solamente por la oportunidad que me ha bridado, sino también por toda la confianza depositada y su apoyo constante. Por ayudarme a reencontrar la motivación cuando esta se escondía y conseguir que me ilusionase con cada idea y proyecto nuevo. Por todo lo que he aprendido de él, pero sobre todo por la gran inspiración que ha sido para mí para llegar a donde a día de hoy estoy. Por todo ello y mucho más, gracias.

A mis compañeros de laboratorio; a María José, por todo lo que he aprendido de ti desde el primer día que puse un pie en el laboratorio, por tu disposición ayudar siempre y tu preocupación contante por el bienestar del grupo. A Carmen Ozuna, por ser mi mentora de laboratorio enseñarme a dar los primeros pasitos en el mundo de la biología molecular. A ambas os agradezco toda la paciencia y ayuda brindada a una novata que empezaba con ilusión en este mundo de la investigación científica. A Ana Adela, por el excelente trabajo técnico que realiza y la inmensa ayuda que supone para todos su colaboración, pero principalmente por su cariño y apoyo. A Lola, por su gran compañerismo y disposición a ayudar siempre con una sonrisa. A Javi, por su ayuda y enseñanzas en mis primeras etapas de formación. A los nuevos fichajes, Helena y Miriam, por todo su apoyo, ayuda y ánimo. También agradecerles a todos los compañeros que han pasado por el laboratorio; Paco, Alba, Ana, Ricardo, Fidel, Pablo, Fortunato, Candelario, Julio y algunos más, con los que he compartido días de trabajo, charlas, ideas y alguna ayuda siempre que fuese posible; y a todos los miembros del departamento de mejora genética vegetal y grupo de doctorandos del Instituto de Agricultura Sostenible.

A Daniel F. Voytas y todo los miembros de su grupo, por recibirme con los brazos abiertos y permitirme completar mi formación en uno de los centros de investigación más prestigiosos de Estados Unidos, en la Universidad de Minnesota, durante mi estancia. Especialmente quiero expresas mi gratitud a Colby Starker, por todo su apoyo brindando y su gran ayuda.

A mi tutor de tesis, Jesús Jorrín, gracias por tu ayuda y por haberme acompañado en esta etapa.

Me gustaría también agradecer a mis amigos, con los que tengo la gran suerte de contar desde hace muchos años, el gran apoyo y ánimo que me han brindado a los largo de estos años, sobre todo por escucharme e interesarse por mi trabajo a pesar de no entender la mitad de las palabras que salían de mi boca, y por preocuparse siempre por el estado y desarrollo de "mis trigos", alegrarse con mis logros y sufrir con mis fracasos. A Marina, por intentar echar una mano siempre hasta en el detalle más pequeño posible, ayudándome a solventar alguna dificultad de edición de imagen. A Carmen, por todos tus mensajes de ánimo, eres el mayor ejemplo para mí de lucha y constancia. A Fabi por su preocupación constante, y animarme siempre que lo necesito a pesar de la distancia. A Alberto, por compartir conmigo de primera mano las penas y alegrías de la vida del doctorando; gracias por tu apoyo, tu ayuda y no dejarme que me pierda ni uno de los plazos de entrega de documentación para el doctorado años tras año. A Estefanía, siempre mostrando un interés especial por mi trabajo, animándome a perseguir mis objetivos y no tirar la toalla. También por vivir conmigo de cerca esta última etapa de nervios en la redacción final de la tesis y por su inmensa y desinteresada ayuda como traductora; puedes ser la persona que más veces se haya leído la introducción de esta tesis y más ganas tenga de que el trigo sin gluten sea una realidad en nuestro plato. Son muchos más los nombres que me gustaría poder mencionar en estas líneas, que por cuestiones de espacio quedan sin aparecer.

A Eu, por ser uno de mis grandes pilares y ser la persona que ha estado siempre a mi lado cuando lo he necesitado y cuando no, por alegrarte y vivir conmigo todos mis logros pero también sufrir conmigo la frustración cuando las cosas que se tuercen. Por ser mi compañero en toda esta etapa, gracias.

Por último, y con especial cariño, quiero agradecerle a mi familia el apoyo incondicional y todo el cariño recibido no solo a lo largo de estos años, sino de toda mi vida. En especial a mis padres, Susana y José Damián, por la educación y los valores que desde pequeña me han inculcado, la formación que me han permitido adquirir y todo el apoyo y el inmenso cariño que siempre he recibido, buscando siempre lo mejor para mí, ellos son los principales responsable de que a día de hoy me halle donde estoy y por supuesto sin ellos esta tesis no hubiese sido hoy posible. Por último y de una manera especial, quiero dedicarle mi trabajo a mi abuela, Susana, quien sé que hubiese vivido con especial ilusión el culmen de mi etapa predoctoral.

Summary

World agricultural production is dominated by three cereal crops; wheat, maize and rice, of which wheat is the most widely distributed, with an annual production of more than 771 million tonnes, only below that of maize. Its wide distribution and use are due to its adaptability to different environments, high yield, and versatility for the food industry where it is used in a wide variety of daily foodstuff such as bread, pasta, cakes, noodles, biscuits, etc. Wheat is also a relevant nutritional source of complex carbohydrates, protein, and dietary fiber, as well as minerals, vitamins and phytoactive compounds, and its consumption is more widespread than that of barley or rye.

Gluten, which is also present in other *Triticeae* species including barley and rye, comprise about 80% of the total wheat grain protein and provides wheat flour with unique technological and biomechanical properties responsible for the bread-making quality of wheat flour. However, gluten is not a single protein, but a complex mixture that include monomeric gliadins (α-, ω- and γ-gliadins) and glutenins (HMW and LMW), which provide, respectively, wheat dough with viscosity and elasticity. The non-gluten proteins (NGP) constitute about the 20% of the total grain protein content and include a wide range of proteins with metabolic and structural functions that plays a secondary role in wheat quality. Within NGP are included various groups of proteins such as α-amylase/trypsin inhibitors (also knowns as ATIs), serine protease inhibitor (serpins), Lipid transfer proteins (LTPs), β-amylase, etc.

Gluten proteins are also responsible of triggering certain pathologies. Gluten-related allergies and intolerances include celiac disease (CD), non-celiac wheat/gluten sensitivity (NCGS) and wheat allergies. CD is usually defined as an autoimmune enteropathy caused by dietary gluten intake in genetically predisposed persons affecting about 1% of western countries population. In contrast, NCGS is not well understood, with symptoms overlapping with those of CD and irritable bowel syndrome (IBS). Prevalence of NCGS is difficult of assess as, in contrast to CD, there is not reliable immunogenic, genetics, or molecular biomarkers, and diagnosis is made by exclusion of other pathologies. It is not clear whether the causal agent of NCGS is gluten; recent studies report that FODMAPs (fermentable oligosaccharides, monosaccharides and disaccharides and polyols) and non-gluten components of wheat such as ATIs seems to activate the innate immune system , and therefore they were proposed to have a role in the NCGS. In the case of CD, there is a vast knowledge about the gluten epitopes that are recognized by HLA-DQ2 and HLA-DQ8 molecules present in predisposed individuals. Gluten peptide recognition by HLA-DQ2/8 molecules stimulates T-cells response which ultimately induces crypt hyperplasia and intestinal villus damage along with several derived complications. Although several CD epitopes are found in the glutenin fraction of gluten, the majority of the immunogenic CD epitopes are found in the gliadin fraction of gluten. Among gliadins, α-gliadins have the

strongest immunogenicity. They contain the 33-mer, which is known to be the main immunodominant peptide for CD as it contains six overlapping copies of three different DQ2-restricted T-cell epitopes with highly stimulatory properties .

A lifelong strict gluten-free diet (GFD) is the only available treatment for CD. In the case of NCGS many patients follow a GFD, although there is evidence that they may have different levels of tolerance. However, a GFD is difficult to follow, socially inconvenient, expensive and tends to be less healthy as they tend to be poorer in several important dietary components with involvement of high amounts of fat and sugar to mimic the viscoelastic properties of gluten. In addition, the GFD may also have negative effects on beneficial gut microbiota.

Given this situation, the development of cereal varieties devoiding of immunogenic epitopes should be considered. Unfortunately, this is almost impossible to achieve through traditional plant breeding due to the high number and complexity of the genes encoding for gluten proteins. Instead, several strategies can be deployed towards this goal, which include: i) the search of varieties that naturally lack of CD epitopes, ii) the development of new cereals with low immunogenic profile and iii) the application of biotechnological technologies to overcome traditional breeding limitations.

Tritordeum is a new cereal derived from the cross between a species of wild barley and durum wheat, which differs from bread wheat in its gluten composition, particularly with respect to gliadin. The reason is the fact that durum wheat lacks the D genome, where the most important epitopes in relation to celiac disease are located, and that barley varieties also contain very few immunogenic peptides compared to wheat, making tritordeum a good candidate as cereal with low levels of gluten immunogenic peptides.

Despite all efforts, no varieties of wheat that naturally lack immunogenic epitopes have been identified so far. The development of biotechnological techniques, such as genetic transformation and interference RNA (RNAi), allowed gene silencing of specific gliadin fractions and all gliadins present in wheat grains where most of relevant epitopes for CD and other gluten pathologies are located. Some of these RNAi lines showed a reduction in their stimulation capacity of T-cell clones, thus showing absence of reactivity and the possibility of manufacturing food for celiac disease patients. However, the silencing of gliadins caused a compensatory effect with other protein fractions, increasing the content of NGP, which is particularly important as proteins belonging to this group, such as ATIs, β-amylase and serpins, have been related to wheat allergens.

On the other hand, in the last decade genetic engineering has expanded its toolbox by incorporating gene editing by the use of sequence-specific nucleases (SSN). CRISPR/Cas9 technology allows precise changes at specific points of interest in the genome. Although it has become one of the most powerful tools for gene editing in plants, examples describing breeding of important traits in crops with complex, polyploid genomes are still limited.

The general objective of this work is the identification and development of healthier and suitable cereals for people suffering from allergies and intolerances. Potential benefits of its development are the improvement of the organoleptic and nutritional profile of foods for CD and NCGS patients, as well as for general population that for whatever reason aims to reduce their gluten intake. With this purpose we characterized of new cereal varieties with reduced immunogenic profile and applied gene editing technologies (CRISPR/Cas9) to target the most immunodominant complex related to CD in wheat.

Chapter one contains a general introduction that gives an overview of main gluten associated allergies and intolerances, allergens, and some strategies for the development of cereals with a low immunogenic profile. Results from this work are described in three chapters, and main conclusions in the last chapter.

Due to the differences in gluten content with wheat, tritordeum is a good candidate as cereal with low levels of gluten immunogenic peptides. **Chapter two** aims to evaluate tritordeum as an alternative cereal for patients suffering NCGS. The response of a group of ten patients diagnosed as NCGS to tritordeum breads was studied to test if they could tolerate this new cereal in their diet. The trial was performed in two phases ("Basal" and "Tritordeum"), each of which lasted seven days. Gastrointestinal symptoms as well as tritordeum acceptability were recorded. Gastrointestinal symptoms of the subjects showed no significant change between the two phases, and tritordeum bread was rated higher than the basal gluten-free bread, particularly in terms of texture and taste. Additionally, we analyses the effect of tritordeum consumption on the intestinal microbiota of these patients, due to the recent idea that a dysbiosis of the intestinal microbiota could contribute to the etiology and pathogenesis of NCGS. The composition and structure of the intestinal microbiota analyzed by sequencing of the 16S showed that there was a trend to increase the relative abundance of some short-chain fatty acids producing bacteria, especially the butyrate-producing genus, *Faecalibacterium*, although the global structure and composition of the intestinal microbiota of each patient was stable after tritordeum bread consumption. This suggests an increase of the intestinal healthy status in these patients, as butyrate contribute to maintenance of the intestinal barrier and is known for its anti-inflammatory effects, promoting the intestinal homeostasis and immune tolerance. Results of this study suggest that tritordeum may be tolerated by at least a sub-set of NCGS sufferers who do not require strict exclusion of gluten from their diet and may benefit from its better organoleptic and nutritional properties.

The stimulating properties of wheat RNAi lines, which differ in the composition of grain proteins (gluten and non-gluten) relevant for CD and other gluten pathologies, were evaluated in **chapter three**. A comparative study of the stimulatory potential of seven low-gluten RNAi lines was carried out through a proteomic analysis by reversed-phase high-performance liquid chromatography and liquid chromatography-tandem mass spectrometry analysis. The downregulation by RNAi of prolamin fractions provided wheat lines differing in the protein composition and in the content of CD immunogenic epitopes.

These RNAi lines showed a reduction in total gliadin content, in specific gliadin fractions (particularly α-gliadins for some lines), and an increment in the HMW, and in the NGP such as ATIs, serpins and triticins. Also, Peripheral Blood Mononuclear Cells (PBMCs) of 35 patients with active CD were used in this study to assess stimulatory response induced by proteins extract from the RNAi lines. Analysis of the proliferative response and INF-γ release of PBMCs demonstrated impaired stimulation in response to all RNAi lines. Stimulatory response was particularly low for three lines, whose response does not significantly differ from that of gluten-free rice protein extract used as a negative control. These lines present a very low content of gluten, as measured by G12 monoclonal antibody, with a pronounced decrease in α-gliadins, and differences in the glutenin content. The NGP seems not to play a key role in PBMCs proliferation and INF-γ release.

In **chapter four** CRISPR/Cas9 technology is applied to target the α-gliadin gene family, where the 33-mer is located. Estimates for α-gliadin gene copy number range from 25–35 to perhaps even 100 copies per haploid genome, which along with the complexity of these genes makes the α-gliadin family especially challenging for genome editing. We designed two sgRNAs (sgAlpha-1 and sgAlpha-2) to target a conserved region adjacent to the coding sequence for the 33-mer in the α-gliadin genes. Twenty-one (15 bread wheat and 6 durum wheat) mutant lines were generated, all showing strong reduction in α-gliadin content as well as other groups of gliadins. High-throughput sequencing demonstrated the efficiency of the CRISPR/Cas9 system to produce mutations in the α-gliadin gene family. Up to 35 different genes were mutated in one of the lines of the 45 different genes identified in the wild type; while immunoreactivity was reduced up to 85% as revealed the R5 and G12 ELISA tests. Transgene-free lines were identified, and no off-target mutations have been detected in any of the potential targets. The low-gluten, transgene-free wheat lines described here constitute an unprecedented advance, and the resultant lines could be used to produce low gluten foodstuff and serve as source material in plant breeding programs to introgress this trait into elite wheat varieties.

Resumen

La producción agrícola mundial está dominada por tres cultivos de cereales: el trigo, el maíz y el arroz, de los cuales el trigo es el más ampliamente distribuido. La producción anual de trigo supera los 771 millones de toneladas (FAOSTAT, 2017), hallándose únicamente por debajo de la del maíz. Su amplia distribución y uso se debe a su gran adaptabilidad a diferentes entornos, su alto rendimiento, y a su versatilidad para la industria alimentaria, en al que se emplea para la fabricación de una gran variedad de alimentos de uso diario como pan, pastas, pasteles, fideos, galletas, etc. El trigo es también una fuente nutricional importante de carbohidratos complejos, proteínas y fibra dietética, así como de minerales, vitaminas y compuestos fitoactivos.

El gluten, también presente en otras especies de la tribu *Triticeae* como la cebada y el centeno, representa alrededor del 80% de la proteína total del grano de trigo y confiere unas propiedades tecnológicas y biomecánicas únicas a la harina de trigo, las cuales son responsables de la calidad harino-panadera del trigo harinero. Sin embargo, el gluten no es una única proteína, sino una mezcla compleja de ellas, donde se incluyen las gliadinas (α-, ω- y γ-gliadinas), monoméricas, y gluteninas (HMW y LMW), poliméricas, que proporcionan viscosidad y elasticidad a la masa de trigo respectivamente. Las proteínas no pertenecientes al gluten (NGP) constituyen alrededor del 20% del contenido total de la proteína del grano e incluyen una amplia gama de proteínas con funciones metabólicas y estructurales que juegan un papel secundario en la calidad del trigo. Dentro de las NGP se incluyen varios grupos de proteínas como α-amilasa/tripsina inhibidores (también conocidos como ATIs), inhibidor de serina proteasa (serpinas), proteínas de transferencia de lípidos (LTPs), β-amilasas, etc.

Las proteínas del gluten también son responsables de desencadenar ciertas patologías. Las alergias e intolerancias relacionadas con el gluten incluyen la enfermedad celíaca (EC), la sensibilidad no celíaca al gluten/trigo (NCGS) y las alergias al trigo. La EC se define como una enteropatía autoinmune causada por la ingesta de gluten que se da en personas genéticamente predispuestas, afectando aproximadamente al 1% de la población de los países occidentales. Por el contrario, la NCGS no está bien definida, con síntomas que se superponen con los del EC y el síndrome del intestino irritable (SII). La prevalencia de la NCGS es difícil establecer debido a la falta de biomarcadores, realizándose su diagnóstico mediante la exclusión de otras patologías. No está totalmente claro si el agente causal de la NCGS es el gluten; estudios recientes han demostrado que los FODMAPs (oligosacáridos, monosacáridos, disacáridos y polioles fermentables) y algunas NGP del trigo, como los ATIs, parecen activar el sistema inmunitario innato y, por lo tanto, desempeñan cierto papel en la NCGS. En el caso de la EC, existe un amplio conocimiento sobre los epítopos del gluten reconocidos por las moléculas HLA-DQ2 y HLA-DQ8 presentes en los individuos genéticamente predispuestos. El reconocimiento de ciertos péptidos por las moléculas HLA-

DQ2/8 estimula la respuesta de las células-T, lo que en última instancia induce a una hiperplasia de las criptas y daños en las vellosidades intestinales además de varias complicaciones derivadas. Aunque existen epítopos de la EC en la fracción de las glutenina, la inmensa mayoría se encuentran en la fracción de las gliadinas, particularmente en las α-gliadinas que son las más inmunogénicas. En ellas se encuentra el 33-mer, conocido por ser el mayor complejo inmunogénico de la EC ya que contiene seis copias superpuestas de tres epítopos con propiedades altamente estimulatorias.

Actualmente el único tratamiento para la EC es una estricta dieta sin gluten (DSG) mantenida de por vida. Aunque la DSG es también recomendada para los pacientes con NCGS, existen evidencias de que puede haber distintos niveles de tolerancia para esos pacientes. Una DSG es difícil de seguir, socialmente inconveniente, costosa económicamente y además tiende a ser menos saludable, ya que es carece de varios componentes nutricionales importantes además de contener mayores cantidades de grasas y azúcares con el objeto de imitar las propiedades viscoelásticas del gluten. Del mismo modo, la DSG también se asocia a efectos negativos sobre el microbioma intestinal beneficioso.

El desarrollo de cereales desprovistos de epítopos inmunogénicos es una posibilidad muy atractiva. Desafortunadamente este objetivo es casi imposible de alcanzar mediante mejora genética vegetal clásica, debido al alto número de copias de los genes que codifican para las proteínas del gluten y su elevada complejidad. En cualquier caso, diversas estrategias pueden aplicarse para la consecución de este objetivo, entre las que se incluyen: i) la búsqueda de variedades que de manera natural carezcan de epítopos de la EC y otros alérgenos, ii) el desarrollo de nuevos cereales con bajo perfil inmunogénico y iii) la aplicación de técnicas biotecnológicas para superar las limitaciones de la mejora clásica.

El tritordeum es un nuevo cereal derivado del cruce entre una especie de cebada silvestre y un trigo duro, que difiere del trigo harinero en la composición del gluten, especialmente en lo que respecta a las gliadinas. Esto se debe a que el trigo duro carece del genoma D, donde se localizan los epítopos más importantes en relación con la enfermedad celíaca, y a que las variedades de cebada también contienen muy pocos péptidos inmunogénicos en comparación con el trigo, lo que convierte al tritordeum en un buen candidato como cereal con bajos niveles de péptidos inmunogénicos.

A pesar de todos los esfuerzos, hasta el momento no se han identificado variedades de trigo que carezcan naturalmente de epítopos inmunogénicos. El desarrollo de la biotecnología, con técnicas como la transformación genética y el ARN de interferencia (ARNi), ha permitido silenciar los genes de fracciones específicas de gliadinas, además de los de todas las gliadinas, presentes en el grano de trigo. Algunas de las líneas de trigo ARNi mostraron una reducción en su capacidad estimulatoria de líneas de células-T específicas para la EC, mostrando así una ausencia de reactividad y la posibilidad ser empleadas en la fabricación de alimentos para pacientes celíacos. Sin embargo, el silenciamiento de las gliadinas provocó un efecto compensatorio con otras fracciones proteicas, aumentando el contenido

de NGP. Esto último es particularmente importante debido a que determinadas proteínas pertenecientes a este grupo, como los ATIs, β-amilasas y serpinas, se han relacionado con los alérgenos del trigo.

A lo largo de la última década el abanico de herramientas para la ingeniería genética se ha visto ampliado gracias a la incorporación de técnicas de edición de genes a través de nucleasas específicas de secuencia (SSN). Estas técnicas, entre las que se incluye CRISPR/Cas9, permiten cambios muy precisos en el genoma y se han convertido en una de las herramientas más poderosas para la edición de genes en las plantas. Sin embargo, los ejemplos de edición de caracteres importantes en cultivos con genomas poliploides complejos son todavía muy limitados.

El objetivo general de este trabajo es la identificación y el desarrollo de cereales más saludables y que puedan ser aptos para las personas que sufren de alergias e intolerancias. Estos cereales tendrían indudables beneficios como la mejora del perfil organoléptico y nutricional de los alimentos para los pacientes con EC y NCGS, así como para la población en general que por cualquier motivo pretenda reducir la ingesta de gluten. Con este propósito hemos abordado la identificación y caracterización de nuevas variedades de cereales con perfil inmunogénico reducido y hemos aplicado las tecnologías de edición genética (CRISPR/Cas9) para editar el mayor complejo inmunogénico relacionado con la enfermedad celíaca.

El presente trabajo se estructura en cinco capítulos; el **primer capítulo** ofrece una visión general de las principales alergias e intolerancias asociadas a los cereales, los alérgenos asociados y algunas estrategias para el desarrollo de cereales con un perfil inmunogénico bajo. Los resultados de este trabajo se describen a lo largo de los tres capítulos siguientes, y por ultimo las principales conclusiones están recogidas en el capítulo cinco.

Por sus diferencias respecto al contenido en gluten con el trigo, el tritordeum es un buen cereal alternativo con bajos niveles de péptidos inmunogénicos. El **capítulo dos** tiene como objetivo evaluar el tritordeum en pacientes que sufren de NCGS, para lo cual se ha comparado el consumo de panes de tritordeum y panes sin gluten, que los pacientes consumen habitualmente, en un grupo de diez pacientes diagnosticados con NCGS. El ensayo se realizó en dos fases ("Fase Basal" y "Fase Tritordeum"), cada una de las cuales duró siete días. Se registraron tanto los síntomas gastrointestinales como la aceptación del pan de tritordeum en comparación con el pan sin gluten. No se presentaron cambios significativos en cuanto a los síntomas gastrointestinales de los sujetos entre las dos fases, y el pan de tritordeum tuvo mejor calificación que el pan sin gluten en cuanto a su aceptación, particularmente en términos de textura y sabor. Adicionalmente, se analizó el efecto del consumo de tritordeum en la microbiota intestinal de estos pacientes, debido a la reciente idea de que una disbiosis de la microbiota intestinal puede contribuir a la etiología y patogénesis de la NCGS. La composición y estructura de la microbiota intestinal se analizó mediante secuenciación del 16S, mostrando que, aunque la estructura y composición global

de la microbiota intestinal de cada paciente era estable tras el consumo de pan de tritordeum, existe una tendencia al aumento de la abundancia relativa de algunas bacterias productoras de ácidos grasos de cadena corta, especialmente de *Faecalibacterium*. Este género de bacterias se conoce por su función productora de butirato, lo que sugiere beneficios saludables para la salud intestinal, ya que los butiratos contribuyen al mantenimiento de la barrera intestinal promoviendo la homeostasis intestinal y la tolerancia inmune. Los resultados de este estudio sugieren, por tanto, que el tritordeum puede ser tolerado al menos por un subconjunto de pacientes con NCGS que no requieren la exclusión estricta del gluten de su dieta y que así podrían beneficiarse de sus mejores propiedades organolépticas y nutricionales.

El **capítulo tres** contiene una evaluación de las propiedades estimulatorias de varias líneas de trigo ARNi que difieren en la composición de las proteínas del grano (gluten y NGP) relevantes para la EC y otras patologías. Se realizó un estudio comparativo del potencial estimulador de siete líneas ARNi de bajo contenido en gluten a través de un análisis proteómico, mediante cromatografía de líquidos de alta resolución en fase reversa y cromatografía líquida acoplada a espectrometría de masas. Estas líneas ARNi mostraron una reducción en el contenido total de gliadinas, mediante una reducción de sus fracciones específicas (particularmente de α-gliadinas para algunas líneas), además de un incremento en las HMW, y en el contenido de NGP como ATIs, serpinas y triticinas. Este estudio incluye un análisis mediante células mononucleares de sangre periférica (PBMCs) de 35 pacientes con EC activa para evaluar la respuesta estimulatoria inducida por los extractos proteicos de las líneas de ARNi. El análisis de la proliferación celular y la liberación de INF-γ por parte de las PBMCs mostraron que la respuesta estimulatoria fue claramente diferente dependiendo de la línea ARNi. Esta respuesta estimulatoria fue particularmente baja en tres de las líneas, cuya respuesta no fue significativamente distinta a la del arroz usado como control negativo sin gluten. Estas líneas presentan un contenido muy bajo de gluten medido por el anticuerpo monoclonal G12, una pronunciada disminución de las α-gliadinas, mientras que el contenido en gluteninas fue variable entre las tres líneas. Las NGP no parecen jugar un papel relevante en la proliferación de las PBMCs y en la liberación de INF-γ.

En el **capítulo cuatro** se aplica la tecnología CRISPR/Cas9 con el objetivo de editar la familia de las α-gliadinas, donde se encuentra el 33-mer. Las estimaciones del número de copias de los genes de α-gliadinas varían de entre 25-35 hasta incluso 100 copias por genoma haploide, lo que, junto con su gran complejidad, hace que esta familia de genes sea especialmente difícil para la edición genética. Se diseñaron dos ARN guías (sgAlpha-1 y sgAlpha-2) para editar una región conservada de los genes de α-gliadinas adyacente a la secuencia que codifica para el 33-mer. Se generaron 21 líneas mutantes (15 de trigo harinero y seis de trigo duro), las cuales mostraron una fuerte reducción en el contenido de α-gliadinas, así como en otros grupos de gliadinas. El análisis mediante secuenciación masiva demostró la alta eficacia del sistema CRISPR/Cas9 para producir mutaciones en esta

compleja familia génica; hasta 35 genes diferentes, de los 45 identificados en la línea usada como control, fueron editados en una de las líneas; consiguiendo una reducción de la inmunoreactividad de hasta un 85% según los anticuerpos monoclonales R5 y G12. Algunas líneas se identificaron como libres de transgenes, y no se hallaron mutaciones "off-target" en otros genes. Las líneas de trigo con bajo contenido en gluten y libres de transgenes descritas en este trabajo constituyen un avance sin precedentes, pudiéndose emplear para la producción de alimentos bajos en gluten además de servir como material de partida en programas de mejora vegetal para introducir este rasgo en variedades élite de trigo.

CHAPTER 1

General introduction and objectives

General introduction and objectives

Cereals

Over 195000 plants are edible, but very few of them make any significant contribution to the human food supply. The overwhelming majority of global staple foods are cereals (Cordain, 1999). World agricultural production is dominated by three cereal crops, wheat, maize and rice, with combined annual production of more than 2900 million tonnes (FAOSTAT, 2017). Wheat is the most widely distributed of the three cereals, with an annual production of more than 771 million tonnes (FAOSTAT, 2017), only below that of maize (Figure 1).

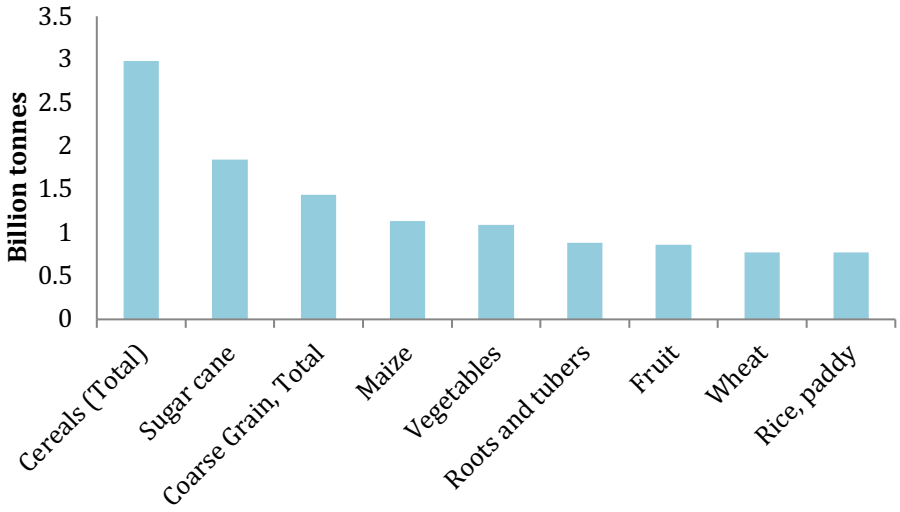


Figure 1. World's Most Produced Commodities (FAOSTAT, 2017)

Cereal crops domestication was originally initiated about 10,000 years ago in the Near East with wheat cultivation. Over time, it extended from the Fertile Crescent to the North and West, eventually reaching Europe, North Africa, and East Asia (Cordain, 1999). This milestone in human history entailed a major shift in the development and evolution of civilizations. Cereal grains consumption enabled the transition from a community of hunter-gatherers and nomadic herders to a more sedentary agrarian-based society where an easier and longer storage of food was feasible. The changing settlement patterns in the population subsequently promoted cultural and technological development. In other words, the development of civilizations cannot be separated from the evolution of cereals, integrating them into culture and even religion (Shewry, 2009)- the key role bread played in the celebration of religious rituals throughout history illustrates it.

The first wheat species to be cultivated were diploid (AA, 2n=2x=14) and tetraploid (AABB, 2n=4x=28), the latter being the result of a cross between the wild ancestor of diploid wheat, *Triticum urartu*, and an unknown species that provided genome the B genome. Current bread wheat (AABBDD, 2n=6x=42) however did not evolve from wild species domestication. Quite on the contrary, it was developed in a human-controlled field by natural hybridization between cultivated tetraploid wheat and a wild species of the genus *Aegilops* (Petersen et al., 2006), providing it with an hexaploide genome by integrating the D genome (Figure 2). Bread wheat flour has unique baking properties, leading it to be favored by farmers, subsequently allowing the evolution of this species. Nowadays, approximately 90% of wheat production is aimed to bread wheat (hexaploid) (Giraldo et al., 2019), whereas the remaining 10% is dedicated to durum wheat (tetraploid), mainly for pasta production. On the contrary, the cultivation of diploid species only occurs in marginal areas.

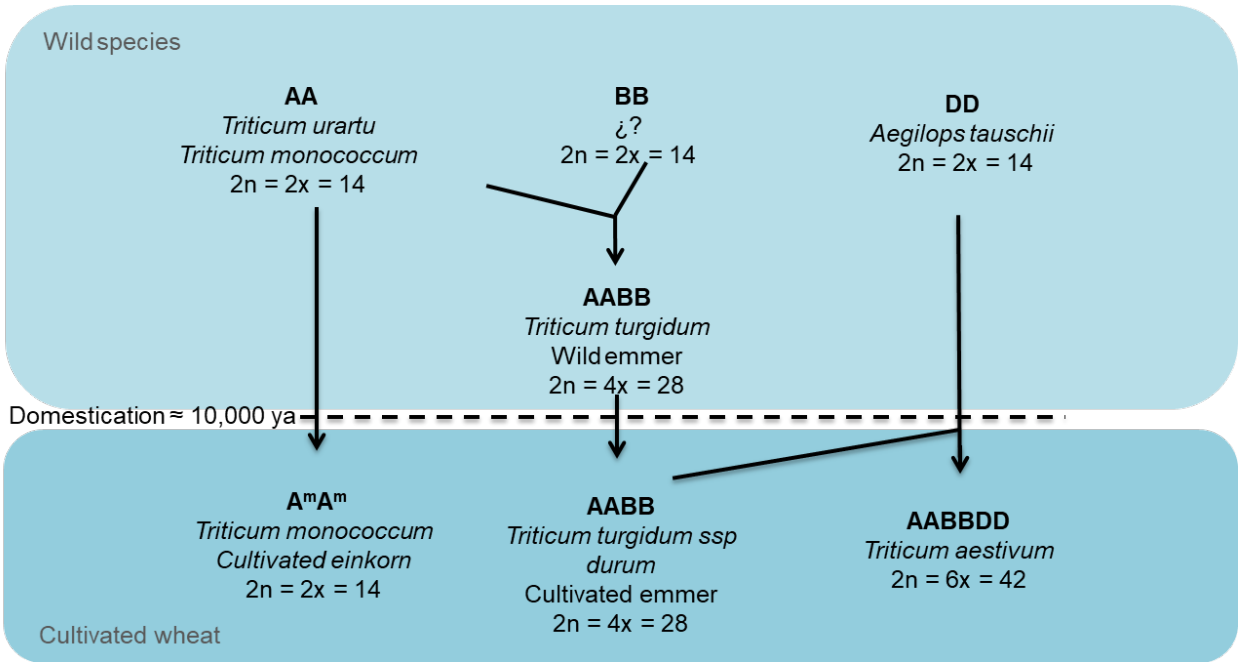


Figure 2. Simplified evolution of modern wheats

Wheat production has gradually expanded worldwide with agricultural revolution, and nowadays wheat-derived foods provide up to 50% of dietary energy intake in both industrialized and developing countries (Tovoli et al., 2015). Wheat is also a relevant nutritional source of complex carbohydrates, protein, and dietary fiber, as well as minerals, vitamins and phytoactive compounds (Shewry and Hey, 2015). It consumption is more widespread than that of barley or rye, and wheat flour versatility allows its usage in a wide variety of daily food products such as bread, pasta, cakes, noodles, biscuits, etc.

Wheat grain is composed mainly of starch, which accounts for 60-75% of its total dry weight. About 80% of the total wheat grain protein corresponds to gluten proteins. This group of proteins, also present in other *Triticeae* tribe species including barley and rye,

provides wheat flour with unique technological and biomechanical properties that play a key role in the dough, ensuring the processing of wheat flour.

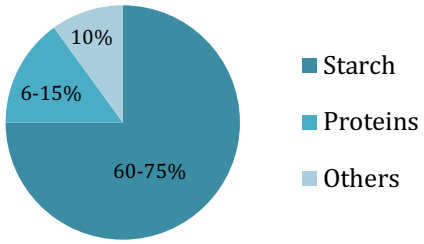


Figure 3. Wheat grain composition (percentage of a dry weight basis)

Allergies and Intolerances to Grains

Cereals and particularly wheat provide exceptional technological and nutritional benefits which make them to be considered as good-quality safe food, as well as to be consumed by billions of people on a daily basis. Nevertheless, the intake of wheat and other related cereals is involved in certain pathologies associated mainly to its protein composition, where gluten proteins play a major role. The main pathologies related with gluten consumption can be classified into three groups (Figure 4) including autoimmune disorders, food and respiratory allergies, and, those pathologies whose mechanism is not yet fully determined, and referred here to as gluten/wheat sensitivity.

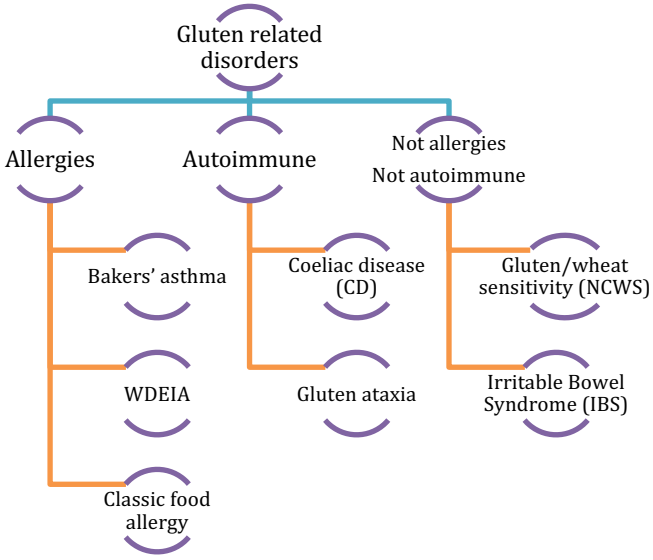


Figure 3. Classification of main gluten/wheat related pathologies. WDEIA, Wheat Dependent Exercise Induced Anaphylaxis.

Autoimmune disorders

There are two autoimmune diseases associated with gluten intake; celiac disease (CD) and gluten ataxia (GA) (Briani et al., 2008). Gluten ataxia causes destruction of the Purkinje neurons of the cerebellum mediated by the immune system, which usually generates symptoms in the motor system. Originally described only in clinical cases, this autoimmune disease is believed to account for up to 40% of idiopathic ataxia diagnoses and, like CD, requires gluten restriction for treatment (Hadjivassiliou et al., 2015).

On the other hand, CD is usually defined as an autoimmune enteropathy caused by dietary gluten intake in genetically predisposed persons (Ludvigsson et al., 2013). This disease presents a complex multifactorial etiology, where gluten intake is a strong environmental component, but there are also genetic and immunological components. CD affects susceptible individuals whose genetic base is located on human chromosome 6, which encodes class II human leukocyte antigens (HLA) and whose most common markers are HLA-DQ2, in 90-95% of patients, and much less frequently HLA-DQ8, present in 5-10% of cases (Sollid, 2002; Tjon et al., 2010). Both markers can be found together in 10% of patients, and similarly, a small proportion of patients (5-10%) are negative for HLA-DQ2 and -DQ8 (Karell et al., 2003), suggesting other factors not yet described.

The most accepted model to explain the immunopathogenesis of CD after gluten ingestion (Figure 5) is a dual model in which both innate and adaptive immune systems are involved, although it is not clear the interactions between both processes (Sollid, 2002; Bernardo and Peña, 2012; Lebwohl et al., 2018). First, there is an innate non-specific immune response that produces a direct effect on the intestinal epithelium, and then a subsequent secondary adaptive response mediated by CD4+ T lymphocytes in the lamina propria. Innate response is induced by certain gluten peptides and it is characterized by enterocytes release of interleukin-15 (IL-15) (Di Sabatino et al., 2006). IL-15 induces NK-response of the intraepithelial lymphocytes, resulting in enterocyte apoptosis and the weakening of the tight junctions in cells not yet entered apoptosis. As a consequence, increases intestinal permeability, allowing gluten peptides to reach the lamina propria and facilitate adaptive immune response.

Two key factors are crucial to the adaptive response and pathogenesis of CD: (i) the resistance of gluten proteins to gastrointestinal digestion, and (ii) the deamidation of glutamine residues present in gluten peptides by tissue transglutaminase 2 (tTG2). There is a great consensus in the characterization of the epitopes present in wheat gluten that are recognized by HLA-DQ2 and HLA-DQ8 and these epitopes are generally rich in proline (P) and glutamine (Q) residues, which makes them resistant to proteolysis and subsequently facilitating their binding to the corresponding HLA molecules present in antigen-presenting cells (APC) (Figure 5). In addition, tissue transglutaminase 2 (tTG2) enzyme found in the gut changes certain residues of glutamine to glutamic acid (a process termed deamidation), making peptides more readily recognizable by HLA molecules. This increased feasibility in

recognition may be due to the introduction of negatively charged residues increasing the affinity of HLA-DQ2 and HLA-DQ8 for gluten peptides (Arentz-Hansen et al., 2000a). Peptides recognition activates an adaptive response mediated by CD4+ T cells, which initiates an inflammatory cascade by releasing cytokines such as interferon-γ (INF- γ), tumor necrosis factor-α (TNF-α) or interleukin-2 (IL-2), as well as metalloproteases and other intermediaries that damage the duodenal and jejunal mucosa inducing crypt hyperplasia and intestinal villus damage. The long-term result is a chronic inflammatory in the small intestine resulting in a flattening of the mucosa. These intestinal lesions are responsible of nutrient malabsorption and lead to derived complications such as anemia, lactose intolerance, amenorrhea, infertility, depression, anxiety, weight loss, growth retardation,

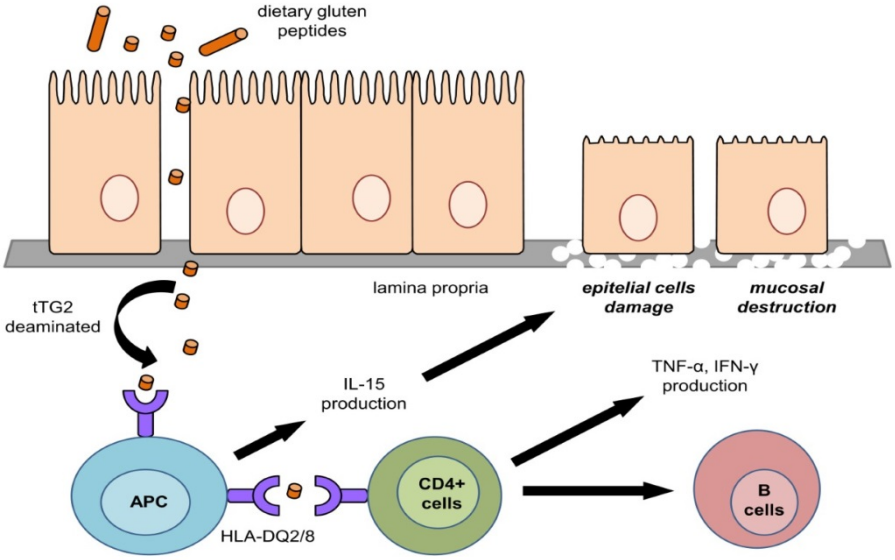


Figure 5. Pathogenesis model for Celiac Disease

elevated liver enzyme levels, and even the appearance of cancer or non-Hodgkin's lymphoma (Green et al., 2001; Catassi et al., 2005; Holtmeier and Caspary, 2006). Gluten ataxia, described above, or dermatitis herpetiformis (Ludvigsson et al., 2013) are considered by some authors as complications derived from celiac disease.

Allergies

Allergic responses to wheat are classified according to the type of immune response that triggers the reaction (Cianferoni, 2016). Both respiratory or food allergy can be controlled by avoiding exposure to wheat or by adding immunomodulatory agents.

In allergies mediated by immunoglobulin-E (IgE), wheat induces symptoms that may include hives, asthma, allergic rhinitis, gastrointestinal distress, and exercise-induced

anaphylaxis (EIA) through the activation of mast cells and basophils in predisposed individuals. Within this field, there is a type of occupational respiratory allergy that occurs in patients with asthma who work in the bakeries, known as "Baker's asthma". For triggering WDEIA (Wheat Dependent Exercise Induced Anaphylaxis) neither wheat nor exercise separately cause the reaction, it is necessary the combination of both factors. It has been suggested that the anaphylactic reaction may be related to the development of large allergenic complexes favored by the activation of transglutaminase in the intestinal mucosa during physical exercise (Palosuo et al., 2003).

On the other hand, in allergies not mediated by IgE, lymphocytic activation induces an inflammatory disorder in which there is an infiltration of eosinophils, cells of the immune system, which invade layers of the stomach and upper gastrointestinal tract, mainly of the duodenum, triggering eosinophilic gastroenteritis (EG) and/or eosinophilic esophagitis (EoE) (Cianferoni, 2016).

Wheat sensitivity

Apart from allergic and autoimmune responses, there are several pathologies not yet fully understood and no consensus exists for their classification. Patients suffering these pathologies develop symptoms overlapping to those of CD after gluten ingestion; however they present an undisturbed duodenal mucosa and absence of CD biomarkers (Elli et al., 2016; Losurdo et al., 2018). This clinical pattern is commonly referred to as non-celiac gluten sensitivity (NCGS), but after evidences that gluten could not be the only dietary component involved in the symptoms (Skodje et al., 2018) non-celiac wheat sensitivity (NCWS) has been proposed as more accurate terminology (Dale et al., 2019).

The typical NCWS symptomatology includes gastrointestinal symptoms such as bloating and abdominal pain, altered bowel habits (diarrhea, constipation or both), nausea and reflux, as well as extraintestinal manifestations such as headache, fatigue, fibromyalgia, anxiety, abdominal pain, disturbed sleep pattern, weight gain, depression, skin rash and dermatitis (Catassi et al., 2013; Peters et al., 2014; Catassi et al., 2015; De Giorgio et al., 2016; Khalid and McMains, 2016). The mechanisms behind this pathology are largely unknown, and a multifactorial process involved in the development of the disease besides an innate immune response has been proposed (Sapone et al., 2010; Catassi et al., 2013; Vazquez-Roque and Oxentenko, 2015). Uhde et al. (2016) also presented evidences of a possible role of intestinal microbiota in these patients given by the translocation of microbial products from the gastrointestinal tract. This fact may contribute to the observed activation of the innate and adaptive immune in NCWS subjects, causing an increase in intestinal permeability (Uhde et al., 2016). However, there are no signs of inflammation caused by gliadins, and there is no activation of basophils or increase in cytokines such as INF-γ (Brottveit et al., 2013; Bucci et al., 2013).

All these factors make diagnosis of NCWS quite complicated, and current diagnosis is often based on the patient's self-diagnosis when there is an improvement of symptoms after excluding gluten from their diet and after exclusion of other pathologies (CD or WA) (Elli et al., 2015). There are no known specific biomarkers for the diagnosis of NCWS and intestinal biopsy is not valid as diagnostic method as no alterations in the intestinal mucosa are observed. Curiously, haplotypes HLA-DQ2 and HLA-DQ8 typically related to CD have been found in up to 50% of diagnosed patients (Sapone et al., 2012). In the absence of a standard protocol, certain recommendations for the diagnosis of NCWS which are included in "The Salerno Experts' Criteria" of NCWS have been proposed. These recommendations include double-blind placebo-controlled gluten challenge, by gluten stimulation after following a gluten-free diet. If there is a variation equal to or greater than 30% in any of the main symptoms, it is considered a positive diagnosis for NCWS (Catassi et al., 2015). Although GFD is recommended for NCWS patients, the gluten tolerance threshold remains unknown, and it has been suggested that individual gluten tolerance levels may vary in affected subjects (Volta et al., 2016). It is not yet fully understood whether lifelong GFD should be recommended nor to what extent other components should also be avoided in the diet (Dale et al., 2019).

The marked nocebo effect observed in gluten challenge is indicative of the existence of other compounds responsible of NCWS symptoms. Although some studies confirms an increase in symptoms in relation to gluten (Bernardo et al., 2007; Di Sabatino et al., 2015), other studies provide important evidence that gluten is neither the only nor the main factor involved. In this sense, a notable improvement in the symptomatology of NCWS patients after reducing the content of FODMAPs (oligosaccharides, disaccharides, monosaccharides and fermentable polyols) in the diet was observed (Biesiekierski et al., 2013; Skodje et al., 2018). However, approximately one third of the patients in the study did show marked changes in symptomatology in response to gluten (Biesiekierski et al., 2013; Zanini et al., 2015). The extraintestinal symptoms reported by NCWS patients are not explained by the known mechanisms of FODMAPs, suggesting that those may be a possible gluten-specific reaction (Peters et al., 2014). Similarly, certain skin conditions reported by a large proportion of NCWS patients improve when they are in a GFD (Bonciolini et al., 2015; Picarelli et al., 2016).

In addition to FODMAPs and gluten, another group of wheat protein knows as ATIs (amylase/trypsin inhibitors) have also been related to the activation of NWGS. These proteins may worsen symptoms in subjects with pre-existing inflammatory diseases, such as autoimmune diseases or chronic inflammation, and several studies show evidence that ATIs are involved in immune response activation and cytokine release, contributing to bowel inflammation and NCWS symptoms (Junker et al., 2012; Catassi et al., 2013; Fasano et al., 2015; Schuppan and Zevallos, 2015). The coexistence of ATIs with gluten in the endosperm of cereals means that a gluten-free diet is equally free of ATIs, so the GFD would effectively reduce the symptomatology of NCWS. Therefore, both ATIs and FODMAPs has

being described as potential triggers of NCWS along with gluten, however further studies are still required to determine exactly their function and mechanism of action in this pathology (Dale et al., 2019).

Another fact that makes the NCWS diagnostic long and complex is that numerous symptoms are very similar to or even overlap with other medical conditions. For example, bloating, abdominal pain and some intestinal irregularities are typical symptoms of Irritable Bowel Syndrome (IBS) (Catassi et al., 2015). FODMAPs are also associated with IBS as they are poorly absorbed in the small intestine and can cause an increase in gas production and light distension, as well as changes in intestinal osmolarity due to fermentation of colon bacteria (Halmos et al., 2014). A diet low in FODMAPs has shown to reduce the symptoms of patients with IBS by up to 80% (Hustoft et al., 2017). As with ATIs, a diet free of gluten-containing foods will automatically lead to a lower intake of FODMAPs, so the response to GFD makes it difficult to distinguish between diagnosis of IBS and NCWS. On the one hand, it was suggested that NCWS is a subgroup of IBS due to the overlapping symptoms, and that the FODMAPs may play a greater role than gluten in symptoms (Vazquez-Roque and Oxentenko, 2015). In contrast, other studies support a specific reaction induced by gluten or wheat in some patients (Capannolo et al., 2015; De Giorgio et al., 2016). However, given that IBS can be assessed as food intolerance seen as gastrointestinal symptoms and NCWS may potentially be an immune response to food antigens that trigger gastrointestinal and extragastrointestinal symptoms, it is likely that IBS and NCWS are different conditions, with overlapping and similar characteristics (Dale et al., 2019).

Prevalence of gluten-related pathologies

Evidence support that gluten-related pathologies have increased in recent years, both in Europe and in the USA, which has led to concern to become a public health problem (LOHI et al., 2007; Rubio–Tapia et al., 2009). The prevalence of CD varies worldwide; it is estimated, based on serological results, between 1.1-1.7% of the world population (Singh et al., 2018). The prevalence of the disease is higher in the North of Europe, being approximately 1-5% in Northern countries. Even so, the highest observed prevalence has been described in an African population living in Western Sahara, with an incidence of 5-6%. Celiac disease is diagnosed in patients of all ages, and approximately 20% are diagnosed after the age of 60. It used to be considered a childhood disease, with a typical symptomatology of weight loss, diarrhea, malabsorption and growth retardation, but this clinical picture is relatively rare nowadays due to a better knowledge and understanding of the disease allowing earlier detection and diagnosis (Dale et al., 2019).

On the other hand, 1-2% of the adult population is affected by wheat allergy; although it is more frequent in children rising up to 2-9% (Dale et al., 2019). This difference between children and adults can be explained by the fact that most patients outgrow their allergy by the age of 16 (Keet et al., 2009). In the case of NCWS the prevalence is difficult to establish

due to the absence of a standardized international diagnosis. According to several authors, the prevalence of NCWS ranges from 0.6% to 13% of the population (Aziz et al., 2016). However, these numbers were calculated based on the results of many self-diagnosed patients, later demonstrated that many of the self-diagnoses met the criteria for IBS rather than NCWS. Therefore, the lack of validated diagnostic criteria, absence of adequate biomarkers, the frequent self-diagnosis, and unconfirmed etiology of reported symptoms, the prevalence of NCWS cannot be accurately established today (Reese et al., 2018).

The increment in the prevalence of gluten-related pathologies cannot be attributed solely to improved diagnostic techniques, and factors may be contributed (LOHI et al., 2007; Rubio-Tapia and Murray, 2010). Causes may include agronomic practices, increased exposure to gluten due to its ubiquity by its use as a food additive, or intense plant breeding towards higher yield but expanding highly immunogenic genotypes (Belderok, 2000; Kasarda, 2013; Ludvigsson et al., 2013). In addition, there is an increased demand for wheat in new markets, especially in countries where wheat was not traditionally grown but a 'western lifestyle' is being adopted, (Ventura et al., 1999; Ivarsson et al., 2000). Other factors such as the expansion of fast industrial processes in bakeries with reduced fermentation times (Gobbetti et al., 2007) have also been proposed as possible factors in the increase of these pathologies.

	Celiac Disease	Wheat allergies	NCWS
Prevalence	1,1-1,7%	1-2%	Unknown
Pathogenesis	Autoimmune reaction	Allergic reaction mediated by Ig-E	Unknown
Antibodies	IgA EMA IgA tTG IgG DGP	IgE for wheat 50% IgA/IgE AGA	None (about 50% have positive IgG AGA)
Enteropathy	Almost always present	Absent	Absent
Symptoms	Intestinal and extra-intestinal	Intestinal and extra-intestinal	Intestinal and extra-intestinal
Complications	Long-term complications, co-morbidities	No co-morbidities, anaphylaxis	Unknown
HLA	95% HLA-DQ2/8	No HLA-DQ2/8 restricted	50% HLA-DQ2/8
Diagnosis	Biopsy Antibodies	Positive IgE for wheat Skin prick test	Excluded WA/CD Effect Gluten challenge
Time of GFD	Lifelong	Possibly lifelong	Unknown

Table 1. Comparison of the main gluten-related pathologies. Modified from *Dale et al., 2019*. NCWS, Non Celiac Wheat Sensitivity.

Wheat proteins: gluten

Wheat is a staple food due to its nutritional properties and technological characteristics. Wheat flour mixed with water produces a network of proteins, linked by inter and intramolecular bonds, which is responsible for the viscoelastic properties and for trapping

of carbon dioxide released during fermentation, providing the typical characteristics of volume and texture of bread and other baked products. Thus, the protein composition of wheat flour has an influence on the properties and functionality of the dough.

Gluten proteins account for up to 80% of the proteins present in the grain. However, gluten is not a single protein, but a complex mixture of them (Figure 6) that accumulates in the endosperm during the development of the grain. This group of proteins is also known as prolamins due to its high content in the amino acids proline (P) and glutamine (Q). Prolamins are further distinguished into two large families of proteins, differentiated on the basis of their solubility, structure, and functionality into gliadins and glutenins (Shewry et al., 1986).

Gliadins are alcohol-soluble proteins, and they are divided into three different types depending on their mobility in electrophoresis of polyacrylamide gels: α-/β-, ω- and γ-gliadins. In addition, there are also small differences between these three subgroups: ω-gliadins have a higher content of glutamine, proline and phenylalanine, making them more polar than the other two groups, while α- and γ-gliadins differ in their proportion of aspartic acid, proline, methionine, tyrosine, phenylalanine and tryptophan (Balakireva and Zamyatnin, 2016). In proportion, α- and γ-gliadins are the majority over ω-gliadins, although the distribution of total gliadins will depend to a large extent on genotype and environmental factors (Wieser and Kieffer, 2001). Gliadins are generally monomeric proteins, contributing to viscosity and extensibility of wheat dough.

The glutenin fraction of gluten is composed of acid or alkali solution soluble proteins (and alcohol insoluble) proteins, which in turn can be subdivided according to their electrophoretic mobility into low molecular weight (LMW) subunits and high molecular weight (HMW) subunits. LMW are the most abundant within the glutenin fraction and they are classified into three groups: groups B, C, D (Payne et al., 1985). The HMW subunits are located on the long arms of the group 1 chromosomes of bread wheat. They are coded by single-copy genes and either type x or y subunits exist tightly linked in each HMW locus, being the x-type (82-90 kDa) larger than the y-type (60-80 kDa) (Shewry et al., 1992). Glutenins form long polymers building an extensive network by disulfide bonds and providing elasticity (HMW), and extensibility and strength (LMW) to wheat dough (Metakovsky et al., 1990; Shewry et al., 2003).

The number of cysteine residues is an important feature as they are involved in inter and intramolecular bonds, determining their structure and properties (Grosch and Wieser, 1999). On this basis, gluten proteins are also classified on their sulfur content (Shewry et al., 1986); the sulfur-rich fractions comprise the α/β- and γ-gliadins, and the LMW, and the sulfur-poor fractions include the ω-gliadins (Figure 6). The HMW subunits are also classified as low-sulfur. Generally α/β- and γ-gliadins contain six and eight cysteine residues respectively, forming intracatenary cross-links (Grosch and Wieser, 1999). However, most of the ω-gliadins cannot establish disulfide bonds due to the lack of cysteine

in their composition. On the other hand, polymeric LMW and HMW form both intra- and interchain bonds, which mediates the aggregation of HMW with the involvement of LMW (Shewry et al., 1986). LMW subunits contain two cysteine residues in N-terminal and C-terminal domains that participate in intermolecular bonds with other gluten proteins, and contributing to the protein network. HMW are the basis of the gluten network backbone formation due to their peculiar structure with at least three and one cysteine residues in their non-repetitive N-terminal and C-terminal domains respectively, flanking a long repetitive domain that typically form β-reverse turns structure. Polymer is established by end-to-end binding of glutenins extending the gluten network where terminators of polymerization were found to be glutathione or gliadins with an odd number of cysteines (Wieser, 2007).

The non-gluten proteins (NGPs) constitute about the 20% of the total grain protein content and include a wide catalogue of proteins with metabolic and structural functions such as components of cell walls, proteins involved in plant defense mechanisms, transport, growth, gene transcription, protein synthesis and various enzymes. NGPs are also divided according to their solubility into two groups, albumins (water-soluble) and globulins (salt-soluble), playing a secondary role in wheat quality.

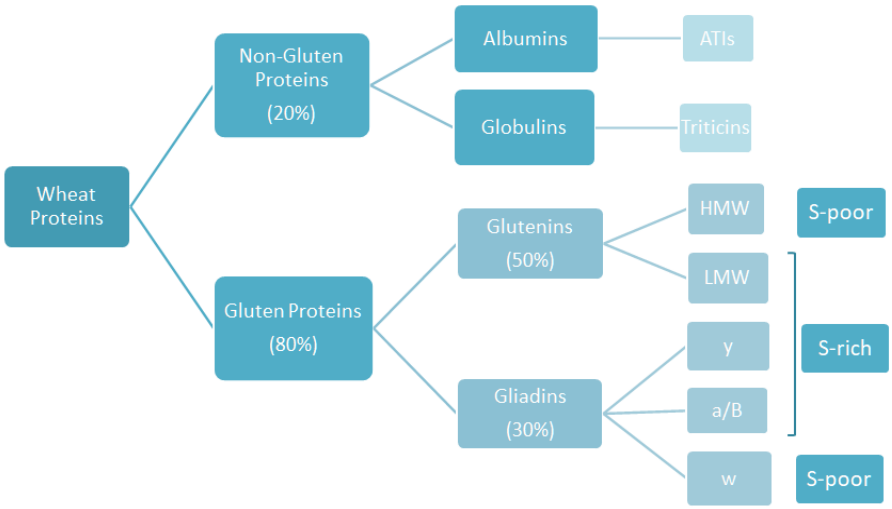


Figure 6. Classification of wheat proteins

Lipid transfer proteins (LTPs), which transfer phospholipids between membranes and vesicles, are part of the NGPs. Non-specific LTPs (nsLTPs) are capable of accommodating various types of lipids and they comprise two subfamilies with molecular weights of approximately 9 and 7 kDa, called nsLTP1 and nsLTP2, respectively (Kalla et al., 1994). It has been observed that nsLTPs, besides their role in lipid trafficking, also participate in plant defense against fungi and bacteria (Hoffmann-Sommergruber, 2002).

Inhibitors of α-amylase constitute between 2-4% of total grain proteins and are also related to defense functions against pathogens (Biesiekierski, 2017). These inhibitors interfere with digestive enzymes in the insect gut, acting as natural pesticides (Franco et al., 2002). Wheat contains two types of amylase inhibitors. One is a bifunctional protein that simultaneously inhibits α-amylase and endogenous subtilisin (serine endopeptidase) from wheat. This protein is called wheat amylase/subtilisin inhibitor or WASI (Mundy et al., 1984). The second group, quantitatively the most important, is known as "CM-like" proteins, due to selective extraction in mixtures of chloroform and methanol, also known as ATIs (α-amylase/trypsin inhibitors). Proteins of this group differ in their spectrum of activity but all of them inhibit the enzymes of mammals and insects without affecting the endogenous amylase of wheat. Other proteins with inhibitory activity among NGPs are the wheat serpins, a serine protease inhibitor identified in almost all organisms (Silverman et al., 2001), acting through irreversible inhibition of proteinases and playing an important role in stress response (Roberts and Hejgaard, 2008).

Cereal allergens

Among wheat flour components, both gluten proteins and non-gluten proteins have been described to cause allergy or intolerance in predisposed individuals. In addition, it has been suggested that some of the non-protein components may trigger certain pathologies related to consumption of cereals (Figure 7).

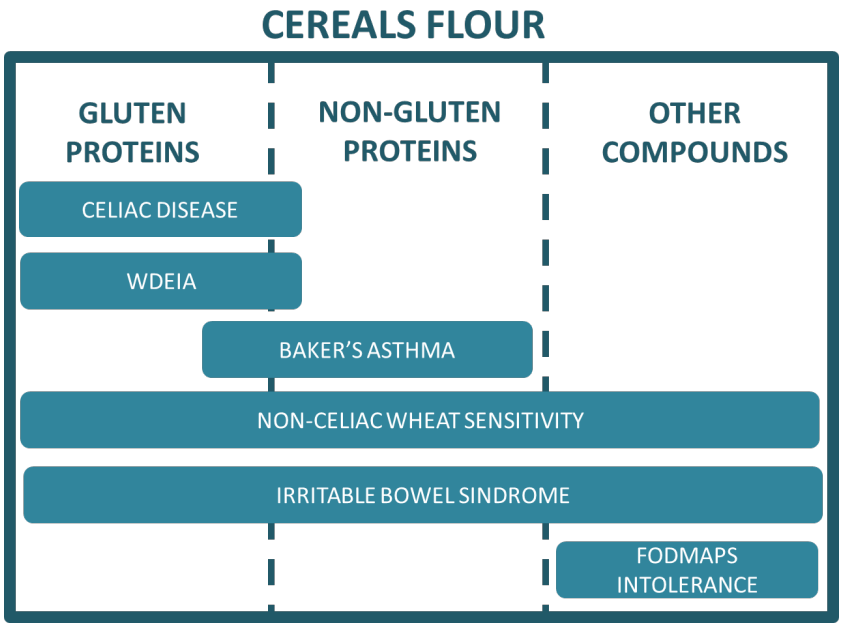


Figure 7. Relationship between the cereal pathologies and the main potentially immunogenic components of flour. WDEIA, Wheat Dependent Exercise Induced Anaphylaxis.

The isolation and characterization of T cells obtained from small intestine biopsies of celiac patients have revealed several epitope sequences recognized by HLA-DQ2 and HLA-DQ8 providing their immunogenic potential. Although several epitopes of CD have been found in glutenins (Molberg et al., 2003), the vast majority (approximately 95%) of immunogenic epitopes have been described in the gliadins fraction (Arentz–Hansen et al., 2002).

The α-gliadins of wheat have the largest number of epitopes (almost 50%) described so far. Using next-generation sequencing (NGS) techniques was possible to differentiate six types of α-gliadins according to the PFPPQQ and PYPQPQ repeat motifs. The different types of gliadins also differ in the number of CD epitopes but, interesting only one type of α-gliadins contains the full set of the 33-mer immunogenic peptide (Ozuna et al., 2015). Moreover, some α-gliadin types do not contain any CD epitopes. The 33-mer peptide was identified as highly immunogenic for CD and contains six overlapping copies of three DQ2 epitopes: one copy of DQ2.5-glia-α1a, two copies of DQ2.5-glia-α1b, and three copies of DQ2.5-glia-α2 (Shan et al., 2002). Other peptides, also relevant for CD are on the α-gliadins; the DQ2.5-glia-α3 (Vader et al., 2002) partially overlapping in the C-terminal of the 33-mer peptide. Additionally, the DQ8-glia-α1 epitope is also present in the C-terminal region of the α-gliadins (van de Wal et al., 1998).

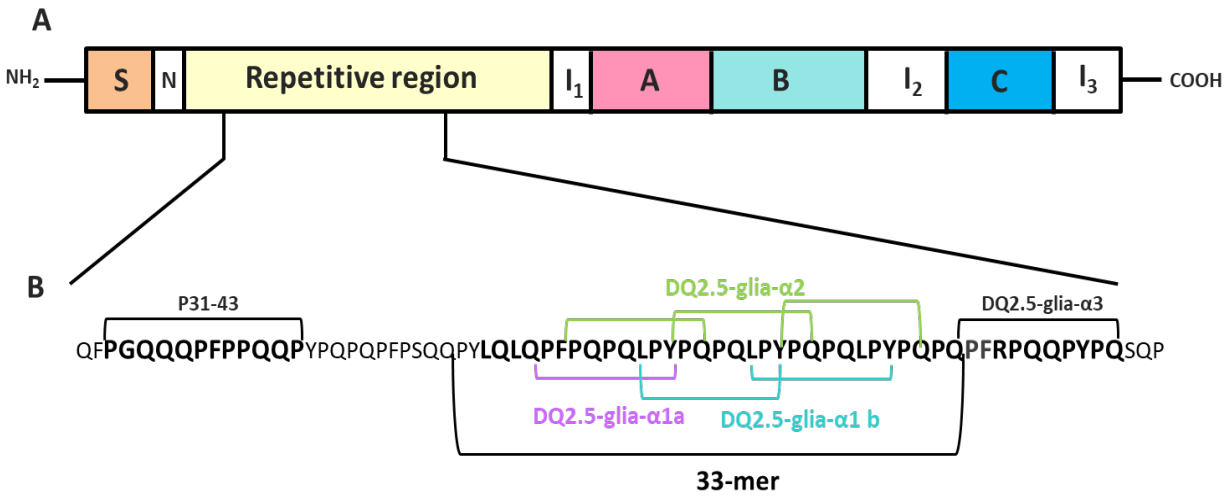


Figure 8. Schematic structure of the α-gliadins (A). Fragment of α-gliadin repetitive region; main immunogenic fragments: peptides p31-43, 33-mer containing six overlapping copies of DQ2 epitopes and DQ2.5-glia-α3 are indicated. (B). A, B, C, conserved regions; S, signal peptide, I_1–I_3, variant regions; N, Nterm.

CD immunogenic epitopes have also been described in γ- and ω-gliadins of wheat, as well as in certain glutenins and in the homolog proteins from other cereals. However, it is not clear the contribution of these immunogenic peptides to the overall immune response in CD. (Anderson et al., 2000; Shan et al., 2002; Tye-Din et al., 2010b; Tanner et al., 2013). Tye-Din et al. (2010b) have thoroughly evaluated more than 16,000 potentially toxic peptides

contained in wheat, barley and rye, and found that only three peptides could elicit most of the immune response observed in CD patients, especially when certain amino acids are deamidated by tTG2. These immunodominant peptides include the 33-mer or any of its variants in the α-gliadins of wheat, a peptide found in the ω-gliadins or homologous proteins in, barley and rye (QPFPQPQQPFPW), and C-hordein from barley (PQQPIPQQPQPYPQQP). Epitopes of relevance to CD have also been described in γ-gliadins, which are recognized by one third of celiac patients (Camarca et al., 2009). There are also peptides related to the activation of the innate immune response in the epithelium and in antigen-presenting cells (APC), required to trigger the adaptive response of T cells (Maiuri et al., 1996; Maiuri et al., 2003). The p31-43 is the most studied peptide activating the innate response to initiate the adaptive response of T cells (Maiuri et al., 2003). This peptide is found in the N-terminal region of α-gliadins (Figure 8). A summary of the main epitopes described for CD so far is given in Table 2.

Epitope name	Prolamin fraction	Sequence	Reference
DQ2.5_glia_α1a	alpha	PFPQPQLPY	(Arentz-Hansen et al., 2000b)
DQ2.5_glia_α1b	alpha	PYPQPQLPY	(Arentz–Hansen et al., 2002)
DQ2.5_glia_α2	alpha	PQPQLPYPQ	(Arentz-Hansen et al., 2000b)
DQ2.5_glia_α3	alpha	FRPQQPYPQ	(Vader et al., 2002)
DQ2.5_glia_γ1	gamma	PQQSFPQQQ	(Sjöström et al., 1998)
DQ2.5_glia_γ2	gamma	IQPQQPAQL	(Vader et al., 2002; Qiao et al., 2005)
DQ2.5_glia_γ2a	gamma	FPQQPQQPF	(Stepniak et al., 2005)
DQ2.5_glia_γ2b	gamma	YPQQPQQPF	(Stepniak et al., 2005)
DQ2.5_glia_γ3	gamma	QQPQQPYPQ	(Arentz–Hansen et al., 2002)
DQ2.5_glia_γ4a	gamma	SQPQQQFPQ	(Arentz–Hansen et al., 2002)
DQ2.5_glia_γ4b	gamma	PQPQQQFPQ	(Qiao et al., 2005)
DQ2.5_glia_γ4c	gamma	QQPQQPFPQ	(Arentz–Hansen et al., 2002)
DQ2.5_glia_γ4d	gamma	PQPQQPFCQ	(Qiao, unpublished)
DQ2.5_glia_γ5	gamma	QQPFPQQPQ	(Arentz–Hansen et al., 2002)
DQ2.5_glia_ω1	omega	PFPQPQQPF	(Tye-Din et al., 2010b)
DQ2.5_glia_ω2	omega	PQPQQPFPW	(Tye-Din et al., 2010b)
DQ2.5_glut_L1	LMW	PFSQQQQPV	(Vader et al., 2002)
DQ2.5_glut_L2	LMW	FSQQQQSPF	(Vader et al., 2002; Stepniak et al., 2005)
DQ2.5_hor_1	hordeins	PFPQPQQPF	(Vader et al., 2003; Tye-Din et al., 2010b)
DQ2.5_hor_2	hordeins	PQPQQPFPQ	(Vader et al., 2003)
DQ2.5_hor_3	hordeins	PIPQQPQPY	(Tye-Din et al., 2010b)

DQ2.5_sec_1	secalins	PFPQPQQPF	(Vader et al., 2003; Tye-Din et al., 2010b)
DQ2.5_sec_2	secalins	PQPQQPFPQ	(Vader et al., 2003)
DQ2.5_ave_1a	avenins	PYPEQQEPF	(Vader et al., 2003; Arentz-Hansen et al., 2004)
DQ2.5_ave_1b	avenins	PYPEQQQPF	(Vader et al., 2003; Arentz-Hansen et al., 2004)
DQ2.2_glut_L1	LMW	PFSQQQQPV	(Bodd et al., 2012)
DQ2.2_glia_α1	alpha	QGSVQPQQL	(Bergseng et al., 2015)
DQ2.2_glia_α2	alpha	QYSQPQQPI	(Bergseng et al., 2015)
DQ8_glia_α1	alpha	QGSFQPSQQ	(van de Wal et al., 1998)
DQ8_glia_γ1a	gamma	QQPQQPFPQ	(Tollefsen et al., 2006a)
DQ8_glia_γ1b	gamma	QQPQQPYPQ	(Tollefsen et al., 2006a)
DQ8_glut_H1	HMW	QGYYPTSPQ	(van de Wal et al., 1999)
DQ8.5_glia_α1	alpha	QGSFQPSQQ	(Kooy-Winkelaar et al., 2011)
DQ8.5_glia_γ1	gamma	PQQSFPQQQ	(Kooy-Winkelaar et al., 2011)
DQ8.5_glut_H1	HMW	QGYYPTSPQ	(Kooy-Winkelaar et al., 2011)
p31-43L	alpha	LGQQQPFPPQQPY	(Maiuri et al., 1996)
p31-43P	alpha	PGQQQPFPPQQPY	(Maiuri et al., 2003)

Table 2. Main CD epitopes described in wheat, rye and barley.

Besides CD epitopes, the ω5-gliadins are also associated with WDEIA (Morita et al., 2003). Most of the WDEIA epitopes consist of 2 glutamine residues flanking 2 or 3 additional residues, generally proline, phenylalanine, serine, glutamic acid, leucine or isoleucine residues, and they may eventually cover up to 45-50% of the sequences of ω5-gliadins (Matsuo et al., 2004; Tatham and Shewry, 2012).

Allergens are also present in NGP, which are associated not only with wheat allergy (Table 3) (Larré et al., 2011; Matsuo et al., 2015), but also as new potential antigens in the humoral response of CD (Huebener et al., 2015). ATIs is an important group of NGP proteins, they are highly resistant to intestinal proteases and have been associated with the release of pro-inflammatory cytokines in both CD and NCWS patients (Sapone et al., 2012; Biesiekierski, 2017). ATIs are also one of the main triggering factors in allergies, such as baker's asthma (Salcedo et al., 2011). Other NGP allergens include β-amylases, LTPs or serpins, which were also categorized as wheat allergens (Tatham and Shewry, 2008; Salcedo et al., 2011). In addition, serpins were also associated as activators of the humoral response in CD (Huebener et al., 2015; Vojdani and Vojdani, 2017).

Allergies	Related allergens
Baker's Asthma	• ATIs • nsLPTs • serpins • β-amylases • gluten proteins
WDEIA	• ω5-gliadins • γ-gliadins • Glutenins

Table 3. Summary of the main wheat allergen proteins related to Baker's Asthma and WDEIA. WDEIA, Wheat Dependent Exercise Induced Anaphylaxis; ATIs, α-amylase/trypsin inhibitors.

Gluten free diet

A strict gluten-free diet (GFD) is the only treatment for gluten-related pathologies (Ventura et al., 1999; Ivarsson et al., 2000). The symptomatology of patients diagnosed with CD or NCWS improves or even disappears when gluten is removed from their diet. However, following a GFD is not that easy, socially inconvenient and it entails a high cost for families.

The exceptional characteristics of gluten proteins have led to an extended usage of gluten as an additive by the food industry, being found in a wide variety of food products that originally do not contain gluten, such as meat, fish, sausages, yoghurt, chocolates, jams and many others. Besides increasing gluten exposure among the population, this results in relatively frequent dietary transgressions among patients who must follow a GFD (Comino et al., 2012).

Currently, wheat, rye, barley and, in some countries, oats are excluded in the production of gluten-free foods. Instead, they are replaced by flour from other cereals (such as maize or rice) and pseudocereals (such as buckwheat, quinoa, etc.), or potato starch. However, products made with these flours tend to be more expensive, often include many additives and they have organoleptic characteristics much different from those made with wheat or other cereals, which worsens their general acceptance (Hager et al., 2012).

Many of the gluten-free products found on the market tend to be less healthy than their gluten analogues since large amounts of fat, sugars and other additives are incorporated in their production to simulate the viscoelastic properties of gluten proteins (Lebwohl et al., 2015; Vaquero et al., 2015; Igbinedion et al., 2017). Therefore, the consumption of industrial gluten-free products is often associated with an excess of saturated and hydrogenated fatty acids, a lack of dietary fiber, a lower daily intake of proteins and an

increase of glycemic index due to the high carbohydrates content (Rosell et al., 2014; Vici et al., 2016). As a consequence, people following a GFD tend to show an unbalanced intake of different nutrients (Gobbetti et al., 2019; Larretxi et al., 2019).

The adherence to a GFD is also associated with a detriment of the intestinal bacterial microbiota, as it deteriorates bacteria populations considered beneficial, such as *Bifidobacterium, B. longum, or Lactobacillus, Clostridium lituseburense*, and *Faecalibacterium prausnitzii*, whilst it increases potentially unhealthy bacterial populations such as *Enterobacteriaceae* and *Escherichia coli* (De Palma et al., 2009). Some studies reported an imbalance in intestinal microbiota (dysbiosis) in CD patients and its contribution to the persistence of some gastrointestinal symptoms. Dysbiosis in NCWS causes intestinal inflammation, diarrhea, constipation, visceral hypersensitivity, abdominal pain, metabolic dysfunction, and problems with peripheral immunity and neuroimmunity (Daulatzai, 2015). Intestinal microbial communities are involved in the well-being of the innate and adaptive immune system, is related to the development of various diseases; hence the importance of maintaining a balanced intestinal microbiota.

Strategies for the development of low-gluten cereals

The development of alternatives to the GFD is an active research area by the scientific community, such as medicine, pharmacology, plant breeding or biotechnology. Several non-dietary therapeutic methods could be considered, which are focused on the development of new drugs to decrease or blocking the appearance of symptoms after the ingestion of gluten. These approaches include the oral use of specific exogenous proteases to eliminate gluten peptides resistant to gastrointestinal proteases (Hausch et al., 2002), the blocking of immunotoxic peptides by gluten-sequestering polymers (Pinier et al., 2012), or the inhibition of tTG2 activity to prevent the deamidation of immunogenic peptides (Badarau et al., 2013). These three strategies, together with the development of other drugs such as desensitizing vaccines, anti-inflammatories or HLA-DQ2/8 blockers, are currently the main pharmacological and medical strategies followed to allow the ingestion of gluten by people with intolerances and allergies (Bonjean et al., 2016).

From the field of plant breeding and biotechnology, this objective is addressed from the perspective of the development of safe cereal varieties suitable for consumption by people suffering from allergies or intolerances. The ideal situation would be to obtain cereal varieties in which the immunogenic components, that triggers the response to gluten, are eliminated but keeping the technological properties and good baking quality. Unfortunately, this is a challenging goal due to the high complexity of the genes encoding for gluten proteins.

Potential benefits of the development of wheat varieties with reduced immunogenic gluten are the improvement of the organoleptic and nutritional profile of foods for CD and NCWS

patients. In addition, general population that for whatever reason wishes to reduce their gluten intake. In fact, a beneficial effect of the low gluten diet compared to high gluten consumption was reported, as the low gluten diet induces changes in the structure and function of the complex intestinal ecosystem of bacteria, reducing hydrogen exhalation, and leading to improvements in self-reported bloating (Hansen et al., 2018).

Natural variability

The identification of cereal varieties that naturally have a low content of immunogenic components has been a very attractive topic in the past years, where considerable efforts have been made. In this sense, numerous studies were reported based on two hypothesis: i) the fact that not all wheat genomes provide to the same amount of immunogenic epitopes, and ii) the proportion of immunogenic gluten between old and modern wheat varieties has changed as consequence of a decrease in the genetic variability associated with intensive agronomic practices towards improved yield but expanding varieties that contain greater number of stimulatory epitopes.

Diploid wheat species have been specially studied to establish the contribution of each of the bread wheat genomes to the total epitopes content; van Herpen et al. (2006) found large differences in the epitope content of diploid wheat species, representing all three genomes of hexaploid wheat, finding highest presence of T cell-stimulatory epitopes in D genome species compared to A and B genome species. This is in agreement with results reported by Spaenij Dekking et al. (2005) who also observed a lower to absent presence of the α-gliadin immunodominant peptides in the A and B genome by studying the presence of epitopes in diploid, tetraploid and hexaploid wheat varieties. Several studies agreed that T. *monococcum* species, AA genome, is the least immunogenic of wheat species (Pizzuti et al., 2006; Gianfrani et al., 2012; Zanini et al., 2015). These studies revealed that some T. *monococcum* lines had a reduced inflammatory effect and they can even be tolerated by some patients, making it a good candidate as a potential use for gluten free foods in combination with other detoxification techniques. In contrast, other studies reported no significant differences between the reactivity of T. *monococcum* in comparison with other polyploid wheat (Šuligoj et al., 2013). The most immunodominant peptide complex for CD, the 33-mer peptide, comes from the D genome of *Ae. Tauschii* (Molberg et al., 2005; Xie et al., 2010; Ozuna et al., 2015), being minority in those coded by genome B (van Herpen et al., 2006). This fact has encouraged the effort for the characterization of tetraploid wheat varieties, lacking the genome D, in the search for a variety with reduced immunogenic properties (Colomba et al., 2012).

Some studies have focused on the comparison between modern and ancient wheat varieties for both hexaploid and tetraploid wheat (Gregorini et al., 2009; van den Broeck et al., 2010) finding a great variability between different species in both, modern and ancient wheats, with some of them showing low content of CD epitopes. However, given the limited number

of trials, it is not possible to conclude that ancient wheat varieties are potentially less immunogenic that all modern ones in reducing chronic disease (Dinu et al., 2018). In fact, Gregorini et al. (2009) found that some ancient varieties such as Graziella Ra and Kamut are more immunogenic that other modern varieties. However, the characterization of grain proteins by HPLC and CD epitopes content in the *Triticeae* has suggested that the process of domestication has contributed to decreasing gluten content in general and in particular the gliadin content and subsequent CD epitopes (Ozuna and Barro, 2018).

Other studies were carried out using wheat lines lacking certain CD epitopes by using nulisomic and chromosomal deletions lines as an alternative to reduce the epitope content. However, lines lacking chromosome 6A showed still provoked a strong immune response (Ciclitira et al., 1980). On the other hand, immunoreactivity tests with lines carrying deletions of the short arm of chromosomes 1D and 6D, where the loci encoding wheat gliadins are found (Figure 9), have shown a significant reduction in the immune response (van den Broeck et al., 2009). Specifically, the deletion of the short arm of chromosome 6D, carrying the Gli-D2 locus (Figure 9) resulted in a significant reduction of T-cell stimulatory capacity (van den Broeck et al., 2009). However, the loss of the short arm of chromosome 6D was also linked to a significant decrease of the technological properties and functionality of the dough.

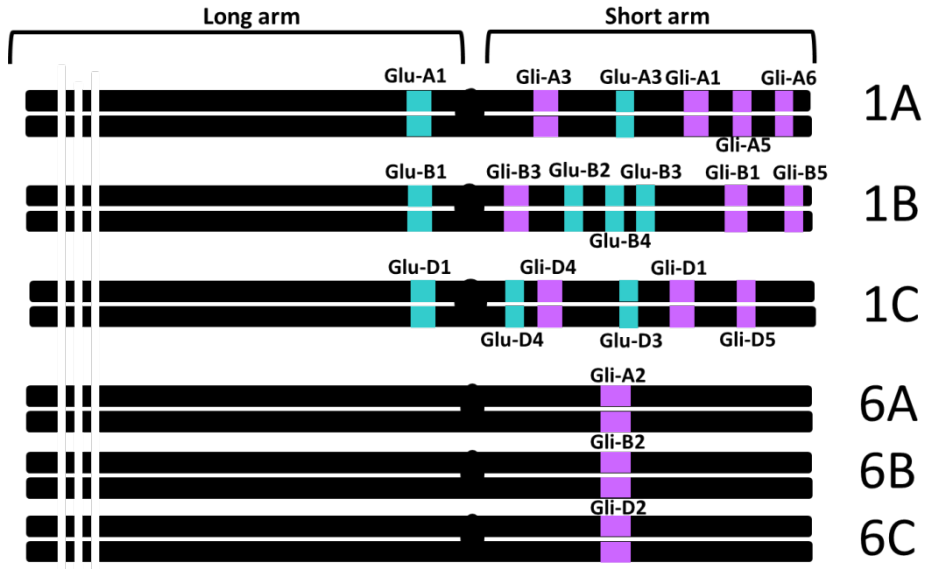


Figure 9. Chromosomal location of the main loci involved in the synthesis of wheat gliadins (Gli) and glutenins (Glu). Genome of origin and chromosome number is indicated on the left.

Several studies also showed that the stimulating capacity of different natural variants of immunogenic peptides or the substitutions of certain amino acids can diminish or even eliminate the immunogenic activity of these CD epitopes (Mitea et al., 2010; Ruiz-Carnicer et al., 2019). Therefore, the identification of gluten proteins lacking T-cell-stimulating epitopes in diploid, tetraploid, and hexaploid wheat varieties would allow the selection of

wheat lines with low reactivity and their incorporation into breeding programs (Spaenij–Dekking et al., 2005; Ozuna et al., 2015).

Development of New Cereals

Triticale was the first synthetic cereal developed in the XIX century, as a result of the hybridization between durum wheat and rye (*Secale cereale*). The resulting cereal combines the yield and quality of wheat with the environmental and disease tolerance of rye. This cereal has not established itself in the food market and its main use is for feeding cattle.

More recently, a new cereal species, named Tritordeum, was produced from the hybridization between a tetraploid wheat (*Triticum durum*) and a species of Chilean wild barley (*Hordeum chilense*). Tritordeum has entered into European market as an alternative to traditional cereals (Martín et al., 1999). The proteins derived from the genome of *H. chilense* improve the strength of tritordeum flour for baking and, in addition, tritordeum contains 5.2 times more carotenoids than durum wheat (Atienza et al., 2007), suggesting a high potential of this species to become a functional food. Tritordeum also has good agronomic yield and interesting physiological traits related to nitrogen metabolism, an essential process in the growth and yield of most crops (Barro et al., 1991).

Tritordeum differs from bread wheat in its gluten protein profile, particularly with respect to the gliadin composition. This difference is due to the fact that durum wheat, the male parental, lacks the D genome, where the most important CD epitopes are located (Molberg et al., 2005), and because barley also contain lower immunogenic peptides compared to bread wheat. These differences in gluten proteins make tritordeum a good candidate as a cereal with low levels of immunogenic peptides (Vaquero et al., 2018). Although it is not suitable for celiac people, it could be of interest for people suffering from other gluten-related pathologies, such as NCWS.

Biotechnological approaches

Biotechnology has emerged in recent decades as an alternative to overcome the limits of traditional breeding. The high complexity of genes encoding for gluten proteins makes almost impossible to use classical plant breeding methods or random mutagenesis to obtain low-gluten varieties. In hexaploid wheat, α-gliadin genes are estimated between 25 and 150 (Anderson et al., 1997), between 17 and 39 for γ-gliadins, and between 15 and 18 for the ω-gliadins (Sabelli and Shewry, 1991). In addition, most of the glutenin genes related to good bread-making properties are linked to gliadins (Anderson et al., 2009). For this reason, new biotechnological tools such as interference RNA (RNAi) or gene editing offer promising tools to obtain wheat varieties with reduced immunogenic epitopes of gluten while maintaining good physical-chemical properties of flour.

RNA of Interference (RNAi)

RNAi is a conserved eukaryotic adaptive response that leads to the specific degradation of target mRNA (Hammond et al., 2000; Bernstein et al., 2001). Initially it was thought to be exclusively a defense mechanism against viruses, but it is also involved in the gene regulation and protection of the genome against transposons. This system is triggers by the presence of double-stranded RNA (dsRNA) that activates a third class ribonuclease known as Dicer, which degrades this dsRNA into small interfering RNA (siRNA) of 21-25 nucleotides in length. The presence of siRNAs activates another enzymatic complex known as RISC (RNA-Induced Silencing Complex), which uses one of the strands of these siRNAs as a template to identify complementary messenger RNA (mRNA), which is degraded, preventing its translation into protein. This mechanism can be adapted for targeted post-transcriptional gene silencing (Figure 10).

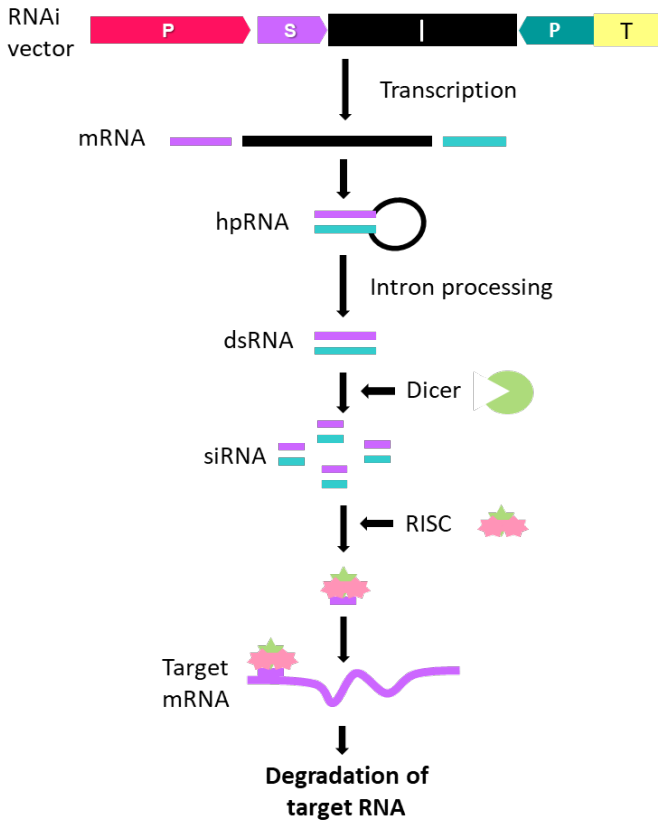


Figure 10. Scheme of RNAi-mediated post-transcriptional down-regulation of a target gene. P, promoter; S, sense inverted repeat; I, intron; A, antisense inverted repeat; T, terminator sequence. *(Source: adapted from Gil-Humanes et al., 2010)*

The RNAi approach has been used to reduce the expression of specific gliadin genes; γ-, ω-, and α-gliadins family (Gil-Humanes et al., 2008; Barro et al., 2016; Altenbach et al., 2019), or even all gliadin genes at once (Gil-Humanes et al., 2010; Wen et al., 2012; Barro et al., 2016). The immunogenic potential of the RNAi lines with all gliadins down-regulated was studied by *in vitro* stimulation of HLA-DQ2 and DQ8 T-cells from celiac patients, resulting in a

marked reduction in proliferative responses of some lines (Gil-Humanes et al., 2010). Additionally, it has been proven that the silencing of gliadins by RNAi activates a compensatory mechanism and the synthesis of other proteins is increased to maintain stable the grain protein content (Pistón et al., 2013; Rosell et al., 2014). For this compensation, the HMW and NGP protein fractions are used (Gil-Humanes et al., 2011; Barro et al., 2016; García-Molina et al., 2017), which is consistent with the fact that some of the RNAi line showed acceptable technological properties for bread making (Gil-Humanes et al., 2014b), and sensory properties similar to breads containing gluten.

Moreover, clinical trials with NCWS patients showed no significant differences, in terms of clinical symptoms and excretion of immunogenic peptides, between the bread made with a RNAi line (named E82) and the gluten-free bread of their choice consumed by volunteers (Haro et al., 2018). Also, an improvement in the intestinal microbiota profile of these patients, with beneficial effects for intestinal permeability was reported after consumption of the low-gluten E82 line (Haro et al., 2018).

Gene editing

In the last decade, genetic engineering has expanded its toolbox by incorporating gene editing through sequence-specific nucleases (SSN) that allows precise changes at specific points of interest in the genome (Samanta et al., 2016). Gene editing allows the production of desired mutations taking advantage of the DNA repair machinery of cells that are activated after a double strand break (DSB) in the DNA. Cells employ two highly conserved pathways to repair DNA (Figure 11); by non-homologous end joining (NHEJ) or by homologous recombination (HR). In NHEJ the DNA is often repaired inaccurately, leading to small insertions or deletions of nucleic acids at the site of rupture. In contrast, HR achieves a precisely DNA repair using a similar sequence as a template.

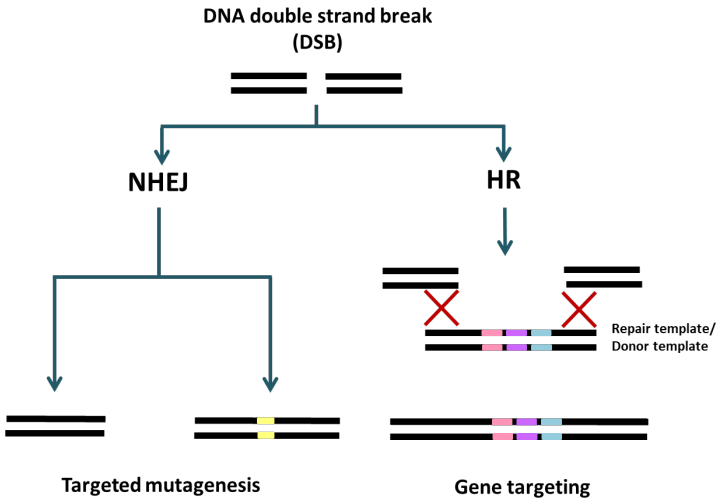


Figure 11. DNA repair mechanism pathways used for Gene Editing

The DSB can be generated at a specific target site through the use of SSNs, which include zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR/Cas) and associated proteins. These SSNs are a combination of a nuclease with a DNA recognition complex that can be either protein (ZFNs and TALENs) or RNA (CRISPR/Cas). The SSNs recognize specific sequences in the DNA and create DSB.

ZFNs consist of a fusion between the catalytic domain of restriction enzyme FokI, and the DNA-binding domain of three or four zinc fingers, where each finger recognizes three nucleotides in the target DNA (Kim et al., 1996). TALENs consist also of fusion to the cleavage domain of the FokI endonuclease with a TALE repeats domain (Christian et al., 2010; Li et al., 2010). Typically, each DNA-binding domain module contains 34 amino acids, and the 12th and 13th residues are the keys to identify the target DNA, known as repeated variable di-residues (RVDs). Both ZFNs and TALENs use FokI endonuclease, and they are designed as a pair of monomers to bind opposite strands of DNA, separated by an appropriate length spacer, allowing FokI monomers to become catalytically active to produce DSBs. Both SSNs require the design, synthesis and delivery of two proteins that act as the DNA-binding domain for each of the sequences to be mutated, and protein engineering is a very complicated, costly, time-consuming and not feasible process in most laboratories.

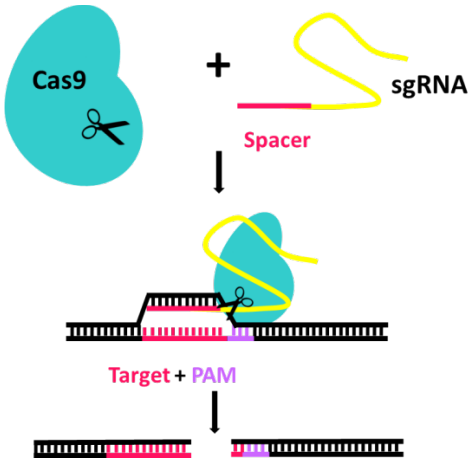


Figure 12. CRISPR-Cas9 target recognition

The CRISPR system was originally described as an adaptive defense of bacteria and archaea against viral infection and adapted for gene editing in plants and animals as a single guide RNA (sgRNA) that retains two critical features (Doudna and Charpentier, 2014); a sequence at the 5' that determines the DNA target site by Watson-Crick base-pairing, and a duplex RNA structure at the 3' that binds to Cas protein (Figure 12). Most CRISPR/Cas systems are based on Cas9 proteins, which recognize a protospacer adjacent motif (PAM) in the DNA

and check for complementarity. When the correct target is found Cas9 binds to DNA and generate a DSB three bases upstream the PAM. Compared to the others SSNs, CRISPR/Cas9 only needs a new sgRNA for each target, while the nuclease remains unmodified. Although CRISPR/Cas9 is less specific than ZFNs or TALENs because the shorter targeting sequence (Cradick et al., 2013), its simplicity and versatility make it the system of choice for gene editing for the precise modification of the genome in many organisms including plants (Table 4) (Sun et al., 2016). In addition, the CRISPR-Cas9 system allows the simultaneous edition of several genes or gene families either by multiplexing the sgRNAs for different target genes or by using conserved regions between them (Xing et al., 2014).

SSNs	Advantages	Disadvantages
ZFN	• Target any genome sequence • Low off-targets mutations	• Time consuming, expensive, laborious, • Target one site each time • Low-efficiency
TALEN	• Target any genome sequence • Low off-targets mutations • Easier design	• Not cost effective and time consuming • Target one site each time • Relatively low-efficiency • Sensitive to target DNA methylation
CRISPR/Cas9	• Simplicity • Target multiple site simultaneously • High efficiency	• Off-target effects • Target selection limited to PAM sequences

Table 4. Comparison between main specific nucleases used in gene editing. ZFN: zinc finger nucleases, TALEN: transcription activator-like effector nuclease, CRISPR/Cas9: clustered regularly interspaced short palindromic repeat Cas9 associated. Source: adapted from *Sun et al. 2016.*

To mediate genome editing in vivo, vector constructs carrying the Cas9 and sgRNA expression cassettes must be delivered into the plant cells. Both expression cassettes can be arranged in separate constructs or single constructs and Agrobacterium-mediated transformation or biobalistic techniques are routine used to integrate CRISPR/Cas9 reagents into plants to produce heritable mutations. To improve Cas9 expression in plants, modified Cas9 genes were optimized with plant-usage bias codons and a single or dual nuclear localization signal (NLS) fused to Cas9 coding sequence to ensure nuclear localization of Cas9. For expression of Cas9 coding sequences in plants, strong constitutive promoter such as those of Ubiquitin gene from maize, rice, and Arabidopsis, and Cauliflower mosaic virus (CaMV) 35S are generally used (Ma et al., 2016). On the other hand, expression of sgRNAs in plants is generally driven by U3 or U6 small nuclear RNA gene promoters, and sgRNAs are transcribed by RNA polymerase III. Other strategies allow the expression of multiple sgRNAs contained in the same expression cassette through the polycistronic tRNA processing system by flanking the sgRNAs with tRNA precursor sequences (Xie et al., 2015). Although CRISPR/Cas9 vectors can be randomly integrated into the plant genome to generate the desired mutations, the transgene can later be eliminated by Mendelian segregation in the offspring.

Objectives

In this context, the objectives of this thesis are:

i) To evaluate tritordeum as an alternative cereal for patients suffering NCWS.

ii) To evaluate the stimulating properties of RNAi wheat lines, which differ in the composition of grain proteins (gluten and non-gluten) to establish optimal protein composition for a maximum reduction of the immunogenic potential.

iii) The application of gene editing technologies (CRISPR/Cas9) to edit the α-gliadins, the most important immunogenic complex related to celiac disease.

The thesis is presented in chapters with the structure commonly used for peer-reviewed scientific publications. At the time of writing, chapters two and three are under review for publication in Nutrients. Chapter three has been published in Plant Biotechnology Journal. Each chapter addresses one of the objectives described in this section.

References

AFSSA (2007) Summary of the Individual and National Study on Food Consumption 2 (INCA 2) 2006-2007.

ALINORM08/31/26 (2008) Codex Alimentarius Commission , Appendix III. Rome, Italy: World Health Organization, WHO.

Altenbach, S.B., Chang, H.-C., Yu, X.B., Seabourn, B.W., Green, P.H., and Alaedini, A. (2019). Elimination of Omega-1,2 Gliadins From Bread Wheat (Triticum aestivum) Flour: Effects on Immunogenic Potential and End-Use Quality. *Frontiers in Plant Science* 10 (580). doi: 10.3389/fpls.2019.00580.

Anderson, O.D., Gu, Y.Q., Kong, X., Lazo, G.R., and Wu, J. (2009). The wheat ω-gliadin genes: structure and EST analysis. *Functional & Integrative Genomics* 9(3), 397-410. doi: 10.1007/s10142-009-0122-2.

Anderson, O.D., Litts, J.C., and Greene, F.C. (1997). The α-gliadin gene family. I. Characterization of ten new wheat α-gliadin genomic clones, evidence for limited sequence conservation of flanking DNA, and Southern analysis of the gene family. *Theoretical and Applied Genetics* 95(1), 50-58. doi: 10.1007/s001220050531.

Anderson, R.P., Degano, P., Godkin, A.J., Jewell, D.P., and Hill, A.V.S. (2000). In vivo antigen challenge in celiac disease identifies a single transglutaminase-modified peptide as the dominant A-gliadin T-cell epitope. *Nature Medicine* 6(3), 337-342. doi: 10.1038/73200.

Arentz-Hansen, H., Fleckenstein, B., Molberg, Ø., Scott, H., Koning, F., Jung, G., et al. (2004). The molecular basis for oat intolerance in patients with celiac disease. *PLoS medicine* 1(1), e1-e1. doi: 10.1371/journal.pmed.0010001.

Arentz-Hansen, H., Körner, R., Molberg, Ø., Quarsten, H., Vader, W., Kooy, Y.M.C., et al. (2000a). The Intestinal T Cell Response to α-Gliadin in Adult Celiac Disease Is Focused on a Single Deamidated Glutamine Targeted by Tissue Transglutaminase. *The Journal of Experimental Medicine* 191(4), 603-612. doi: 10.1084/jem.191.4.603.

Arentz-Hansen, H., Körner, R., Molberg, Ø., Quarsten, H., Vader, W., Kooy, Y.M.C., et al. (2000b). The Intestinal T Cell Response to α-Gliadin in Adult Celiac Disease Is Focused on a Single Deamidated Glutamine Targeted by Tissue Transglutaminase. *The Journal of Experimental Medicine* 191(4), 603. doi: 10.1084/jem.191.4.603.

Arentz–Hansen, H., McAdam, S.N., Molberg, Ø., Fleckenstein, B., Lundin, K.E.A., Jørgensen, T.J.D., et al. (2002). Celiac lesion T cells recognize epitopes that cluster in regions of gliadins rich in proline residues. *Gastroenterology* 123(3), 803-809. doi: 10.1053/gast.2002.35381.

Atienza, S.G., Ballesteros, J., Martín, A., and Hornero-Méndez, D. (2007). Genetic Variability of Carotenoid Concentration and Degree of Esterification among Tritordeum (×Tritordeum Ascherson et Graebner) and Durum Wheat Accessions. *Journal of Agricultural and Food Chemistry* 55(10), 4244-4251. doi: 10.1021/jf070342p.

Aziz, I., Dwivedi, K., and Sanders, D.S. (2016). From coeliac disease to noncoeliac gluten sensitivity; should everyone be gluten free? *Current Opinion in Gastroenterology* 32(2), 120-127. doi: 10.1097/mog.0000000000000248.

Badarau, E., Collighan, R.J., and Griffin, M. (2013). Recent advances in the development of tissue transglutaminase (TG2) inhibitors. *Amino Acids* 44(1), 119-127. doi: 10.1007/s00726-011-1188-4.

Balakireva, A.V., and Zamyatnin, A.A. (2016). Properties of Gluten Intolerance: Gluten Structure, Evolution, Pathogenicity and Detoxification Capabilities. *Nutrients* 8(10), 644.

Barro, F., Fontes, A.G., and Maldonado, J.M. (1991). Organic nitrogen content and nitrate and nitrite reductase activities in tritordeum and wheat grown under nitrate or ammonium. *Plant and Soil* 135(2), 251-256. doi: 10.1007/BF00010913.

Barro, F., Iehisa, J.C.M., Giménez, M.J., García-Molina, M.D., Ozuna, C.V., Comino, I., et al. (2016). Targeting of prolamins by RNAi in bread wheat: effectiveness of seven silencing-fragment combinations for obtaining lines devoid of coeliac disease epitopes from highly immunogenic gliadins. *Plant Biotechnology Journal* 14(3), 986-996. doi: 10.1111/pbi.12455.

Belderok, B. (2000). Developments in bread-making processes. *Plant Foods for Human Nutrition* 55(1), 1-14. doi: 10.1023/A:1008199314267.

Bergseng, E., Dørum, S., Arntzen, M.Ø., Nielsen, M., Nygård, S., Buus, S., et al. (2015). Different binding motifs of the celiac disease-associated HLA molecules DQ2.5, DQ2.2, and DQ7.5 revealed by relative quantitative proteomics of endogenous peptide repertoires. *Immunogenetics* 67(2), 73-84. doi: 10.1007/s00251-014-0819-9.

Bernardo, D., Garrote, J.A., Fernández-Salazar, L., Riestra, S., and Arranz, E. (2007). Is gliadin really safe for non-coeliac individuals? Production of interleukin 15 in biopsy culture from non-coeliac individuals challenged with gliadin peptides. *Gut* 56(6), 889-890. doi: 10.1136/gut.2006.118265.

Bernardo, D., and Peña, A.S. (2012). Developing strategies to improve the quality of life of patients with gluten intolerance in patients with and without coeliac disease. *European Journal of Internal Medicine* 23(1), 6-8. doi: 10.1016/j.ejim.2011.09.016.

Bernstein, E., Denli, A.M., and Hannon, G.J. (2001). The rest is silence. RNA 7(11), 1509-1521.

Biesiekierski, J.R. (2017). What is gluten? *Journal of Gastroenterology and Hepatology* 32(S1), 78-81. doi: 10.1111/jgh.13703.

Biesiekierski, J.R., Peters, S.L., Newnham, E.D., Rosella, O., Muir, J.G., and Gibson, P.R. (2013). No Effects of Gluten in Patients With Self-Reported Non-Celiac Gluten Sensitivity After Dietary Reduction of Fermentable, Poorly Absorbed, Short-Chain Carbohydrates. *Gastroenterology* 145(2), 320-328.e323. doi: 10.1053/j.gastro.2013.04.051.

Bodd, M., Kim, C.Y., Lundin, K.E.A., and Sollid, L.M. (2012). T-Cell Response to Gluten in Patients With HLA-DQ2.2 Reveals Requirement of Peptide-MHC Stability in Celiac Disease. *Gastroenterology* 142(3), 552-561. doi: 10.1053/j.gastro.2011.11.021.

Bonciolini, V., Bianchi, B., Del Bianco, E., Verdelli, A., and Caproni, M. (2015). Cutaneous Manifestations of Non-Celiac Gluten Sensitivity: *Clinical Histological and Immunopathological Features*. Nutrients 7(9), 7798-7805.

Bonjean, A.P., Angus, W.J., and van Ginkel, M. (2016). The World Wheat Book volume 3 : a history of wheat breeding. Paris : Lavoisier.

Briani, C., Samaroo, D., and Alaedini, A. (2008). Celiac disease: From gluten to autoimmunity. *Autoimmunity Reviews* 7(8), 644-650. doi: 10.1016/j.autrev.2008.05.006.

Brottveit, M., Beitnes, A.C., Tollefsen, S., Bratlie, J.E., Jahnsen, F.L., Johansen, F.E., et al. (2013). Mucosal cytokine response after short-term gluten challenge in celiac disease and non-celiac gluten sensitivity. *The American Journal of Gastroenterology* 108(5), 842-850. doi: 10.1038/ajg.2013.91.

Bucci, C., Zingone, F., Russo, I., Morra, I., Tortora, R., Pogna, N., et al. (2013). Gliadin Does Not Induce Mucosal Inflammation or Basophil Activation in Patients With Nonceliac Gluten Sensitivity. *Clinical Gastroenterology and Hepatology* 11(10), 1294-1299. doi: 10.1016/j.cgh.2013.04.022.

Camarca, A., Anderson, R.P., Mamone, G., Fierro, O., Facchiano, A., Costantini, S., et al. (2009). Intestinal T Cell Responses to Gluten Peptides Are Largely Heterogeneous: Implications for a Peptide-Based Therapy in Celiac Disease. *The Journal of Immunology* 182(7), 4158. doi: 10.4049/jimmunol.0803181.

Capannolo, A., Viscido, A., Barkad, M.A., Valerii, G., Ciccone, F., Melideo, D., et al. (2015). Non-Celiac Gluten Sensitivity among Patients Perceiving Gluten-Related Symptoms. *Digestion* 92(1), 8-13. doi: 10.1159/000430090.

Catassi, C., Bai, J.C., Bonaz, B., Bouma, G., Calabrò, A., Carroccio, A., et al. (2013). Non-Celiac Gluten Sensitivity: The New Frontier of Gluten Related Disorders. *Nutrients* 5(10), 3839-3853.

Catassi, C., Bearzi, I., and Holmes, G.K.T. (2005). Association of celiac disease and intestinal lymphomas and other cancers. *Gastroenterology* 128(4,), S79-S86. doi: 10.1053/j.gastro.2005.02.027.

Catassi, C., Elli, L., Bonaz, B., Bouma, G., Carroccio, A., Castillejo, G., et al. (2015). Diagnosis of Non-Celiac Gluten Sensitivity (NCGS): The Salerno Experts' Criteria. *Nutrients* 7(6), 4966-4977. doi: 10.3390/nu7064966.

Cianferoni, A. (2016). Wheat allergy: diagnosis and management. *Journal of asthma and allergy* 9, 13-25. doi: 10.2147/JAA.S81550.

Ciclitira, P.J., Hunter, J.O., and Lennox, E.S. (1980). Clinical testing of bread made from nullisomic 6A wheats in celiac patients. *The Lancet* 316(8188), 234-236. doi: 10.1016/S0140-6736(80)90123-3.

Colomba, M., Vischi, M., and Gregorini, A. (2012). Molecular Characterization and Comparative Analysis of Six Durum Wheat Accessions Including Graziella Ra. *Plant Molecular Biology Reporter* 30(1), 168-175. doi: 10.1007/s11105-011-0328-z.

Comino, I., Real, A., Vivas, S., Síglez, M.Á., Caminero, A., Nistal, E., et al. (2012). Monitoring of gluten-free diet compliance in celiac patients by assessment of gliadin 33-mer equivalent epitopes in feces. *The American Journal of Clinical Nutrition* 95(3), 670-677. doi: 10.3945/ajcn.111.026708.

Cordain, L. (1999). Cereal Grains: Humanity´s Double-Edged Sword. *World Review of Nutrition and Dietetics* 84(1), 19-73. doi: doi.org/10.1159/000059677.

Cradick, T.J., Fine, E.J., Antico, C.J., and Bao, G. (2013). CRISPR/Cas9 systems targeting β-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Research* 41(20), 9584-9592. doi: 10.1093/nar/gkt714.

Christian, M., Cermak, T., Doyle, E.L., Schmidt, C., Zhang, F., Hummel, A., et al. (2010). Targeting DNA Double-Strand Breaks with TAL Effector Nucleases. Genetics 186(2), 757. doi: 10.1534/genetics.110.120717.

Dale, H.F., Biesiekierski, J.R., and Lied, G.A. (2019). Non-coeliac gluten sensitivity and the spectrum of gluten-related disorders: an updated overview. *Nutrition Research Reviews* 32(1), 28-37. doi: 10.1017/S095442241800015X.

Daulatzai, M.A. (2015). Non-Celiac Gluten Sensitivity Triggers Gut Dysbiosis, Neuroinflammation, Gut-Brain Axis Dysfunction, and Vulnerability for Dementia. *CNS and Neurological Disorders Drug Targets* 14(1), 110-131. doi: 10.2174/1871527314666150202152436.

De Giorgio, R., Volta, U., and Gibson, P.R. (2016). Sensitivity to wheat, gluten and FODMAPs in IBS: facts or fiction? *Gut* 65(1), 169. doi: 10.1136/gutjnl-2015-309757.

De Palma, G., Nadal, I., Collado, M.C., and Sanz, Y. (2009). Effects of a gluten-free diet on gut microbiota and immune function in healthy adult human subjects. *British Journal of Nutrition* 102(8), 1154-1160. doi: 10.1017/S0007114509371767.

Di Sabatino, A., Ciccocioppo, R., Cupelli, F., Cinque, B., Millimaggi, D., Clarkson, M.M., et al. (2006). Epithelium derived interleukin 15 regulates intraepithelial lymphocyte Th1 cytokine production, cytotoxicity, and survival in coeliac disease. *Gut* 55(4), 469. doi: 10.1136/gut.2005.068684.

Di Sabatino, A., Volta, U., Salvatore, C., Biancheri, P., Caio, G., De Giorgio, R., et al. (2015). Small Amounts of Gluten in Subjects With Suspected Nonceliac Gluten Sensitivity: A Randomized, Double-Blind, Placebo-Controlled, Cross-Over Trial. *Clinical Gastroenterology and Hepatology* 13(9), 1604-1612.e1603. doi: 10.1016/j.cgh.2015.01.029.

Dinu, M., Whittaker, A., Pagliai, G., Benedettelli, S., and Sofi, F. (2018). Ancient wheat species and human health: Biochemical and clinical implications. *The Journal of Nutritional Biochemistry* 52, 1-9. doi: 10.1016/j.jnutbio.2017.09.001.

Doudna, J.A., and Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. *Science* 346(6213), 1258096. doi: 10.1126/science.1258096.

Elli, L., Branchi, F., Tomba, C., Villalta, D., Norsa, L., Ferretti, F., et al. (2015). Diagnosis of gluten related disorders: Celiac disease, wheat allergy and non-celiac gluten sensitivity. *World journal of gastroenterology* 21(23), 7110-7119. doi: 10.3748/wjg.v21.i23.7110.

Elli, L., Tomba, C., Branchi, F., Roncoroni, L., Lombardo, V., Bardella, M.T., et al. (2016). Evidence for the Presence of Non-Celiac Gluten Sensitivity in Patients with Functional Gastrointestinal Symptoms: Results from a Multicenter Randomized Double-Blind Placebo-Controlled Gluten Challenge. *Nutrients* 8(2), 84.

Fasano, A., Sapone, A., Zevallos, V., and Schuppan, D. (2015). Nonceliac Gluten Sensitivity. *Gastroenterology* 148(6), 1195-1204. doi: 10.1053/j.gastro.2014.12.049.

Franco, O.L., Rigden, D.J., Melo, F.R., and Grossi-de-Sá, M.F. (2002). Plant α-amylase inhibitors and their interaction with insect α-amylases. *European Journal of Biochemistry* 269(2), 397-412. doi: 10.1046/j.0014-2956.2001.02656.x.

García-Molina, M.D., Muccilli, V., Saletti, R., Foti, S., Masci, S., and Barro, F. (2017). Comparative proteomic analysis of two transgenic low-gliadin wheat lines and non-transgenic wheat control. *Journal of Proteomics* 165, 102-112. doi: 10.1016/j.jprot.2017.06.010.

Gianfrani, C., Maglio, M., Rotondi Aufiero, V., Camarca, A., Vocca, I., Iaquinto, G., et al. (2012). Immunogenicity of monococcum wheat in celiac patients. *The American Journal of Clinical Nutrition* 96(6), 1339-1345. doi: 10.3945/ajcn.112.040485.

Gil-Humanes, J., Pistón, F., Barro, F., and Rosell, C.M. (2014b). The Shutdown of Celiac Disease-Related Gliadin Epitopes in Bread Wheat by RNAi Provides Flours with Increased Stability and Better Tolerance to Over-Mixing. *PLOS ONE* 9(3), e91931. doi: 10.1371/journal.pone.0091931.

Gil-Humanes, J., Pistón, F., Hernando, A., Alvarez, J.B., Shewry, P.R., and Barro, F. (2008). Silencing of γ-gliadins by RNA interference (RNAi) in bread wheat. *Journal of Cereal Science* 48(3), 565-568. doi: 10.1016/j.jcs.2008.03.005.

Gil-Humanes, J., Pistón, F., Shewry, P.R., Tosi, P., and Barro, F. (2011). Suppression of gliadins results in altered protein body morphology in wheat. *Journal of Experimental Botany* 62(12), 4203-4213. doi: 10.1093/jxb/err119.

Gil-Humanes, J., Pistón, F., Tollefsen, S., Sollid, L.M., and Barro, F. (2010). Effective shutdown in the expression of celiac disease-related wheat gliadin T-cell epitopes by RNA interference. *PNAS* 107(39), 7023–17028. doi: 10.1073/pnas.1007773107.

Giraldo, P., Benavente, E., Manzano-Agugliaro, F., and Gimenez, E. (2019). Worldwide Research Trends on Wheat and Barley: A Bibliometric Comparative Analysis. *Agronomy* 9(7), 352.

Gobbetti, M., De Angelis, M., Di Cagno, R., Calasso, M., Archetti, G., and Rizzello, C.G. (2019). Novel insights on the functional/nutritional features of the sourdough fermentation. *International Journal of Food Microbiology* 302, 103-113. doi: 10.1016/j.ijfoodmicro.2018.05.018.

Gobbetti, M., Giuseppe Rizzello, C., Di Cagno, R., and De Angelis, M. (2007). Sourdough lactobacilli and celiac disease. *Food Microbiology* 24(2), 187-196. doi: 10.1016/j.fm.2006.07.014.

Green, P.H.R., Stavropoulos, S.N., Panagi, S.G., Goldstein, S.L., McMahon, D.J., Absan, H., et al. (2001). Characteristics of Adult Celiac Disease in The Usa: Results of A National Survey. *American Journal of Gastroenterology* 96(1), 126-131. doi: 10.1111/j.1572-0241.2001.03462.x.

Gregorini, A., Colomba, M., Ellis, H.J., and Ciclitira, P.J. (2009). Immunogenicity Characterization of Two Ancient Wheat α-Gliadin Peptides Related to Coeliac Disease. *Nutrients* 1(2), 276-290.

Grosch, W., and Wieser, H. (1999). Redox Reactions in Wheat Dough as Affected by Ascorbic Acid. *Journal of Cereal Science* 29(1), 1-16. doi: 10.1006/jcrs.1998.0218.

Hadjivassiliou, M., Sanders, D.D., and Aeschlimann, D.P. (2015). Gluten-Related Disorders: Gluten Ataxia. *Digestive Diseases* 33(2), 264-268. doi: 10.1159/000369509.

Hager, A.-S., Wolter, A., Jacob, F., Zannini, E., and Arendt, E.K. (2012). Nutritional properties and ultra-structure of commercial gluten free flours from different botanical sources compared to wheat flours. *Journal of Cereal Science* 56(2), 239-247. doi: 10.1016/j.jcs.2012.06.005.

Halmos, E.P., Power, V.A., Shepherd, S.J., Gibson, P.R., and Muir, J.G. (2014). A Diet Low in FODMAPs Reduces Symptoms of Irritable Bowel Syndrome. *Gastroenterology* 146(1), 67-75.e65. doi: 10.1053/j.gastro.2013.09.046.

Hammond, S.M., Bernstein, E., Beach, D., and Hannon, G.J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *Nature* 404(6775), 293-296. doi: 10.1038/35005107.

Hansen, L.B.S., Roager, H.M., Søndertoft, N.B., Gøbel, R.J., Kristensen, M., Vallès-Colomer, M., et al. (2018). A low-gluten diet induces changes in the intestinal microbiome of healthy Danish adults. *Nature Communications* 9(1), 4630. doi: 10.1038/s41467-018-07019-x.

Haro, C., Villatoro, M., Vaquero, L., Pastor, J., Giménez, M.J., Ozuna, C.V., et al. (2018). The Dietary Intervention of Transgenic Low-Gliadin Wheat Bread in Patients with Non-Celiac Gluten Sensitivity (NCGS) Showed No Differences with Gluten Free Diet (GFD) but Provides Better Gut Microbiota Profile. *Nutrients* 10(12), 1964.

Hausch, F., Shan, L., Santiago, N.A., Gray, G.M., and Khosla, C. (2002). Intestinal digestive resistance of immunodominant gliadin peptides. *American Journal of Physiology-Gastrointestinal and Liver Physiology* 283(4), G996-G1003. doi: 10.1152/ajpgi.00136.2002.

Hoffmann-Sommergruber, K. (2002). Pathogenesis-related (PR)-proteins identified as allergens. *Biochemical Society Transactions* 30(6), 930. doi: 10.1042/bst0300930.

Holtmeier, W., and Caspary, W.F. (2006). Celiac disease. Orphanet J Rare Dis 1, 3. doi: 10.1186/1750-1172-1-3.

Huebener, S., Tanaka, C.K., Uhde, M., Zone, J.J., Vensel, W.H., Kasarda, D.D., et al. (2015). Specific Nongluten Proteins of Wheat Are Novel Target Antigens in Celiac Disease Humoral Response. *Journal of Proteome Research* 14(1), 503-511. doi: 10.1021/pr500809b.

Hustoft, T.N., Hausken, T., Ystad, S.O., Valeur, J., Brokstad, K., Hatlebakk, J.G., et al. (2017). Effects of varying dietary content of fermentable short-chain carbohydrates on symptoms, fecal microenvironment, and cytokine profiles in patients with irritable bowel syndrome. *Neurogastroenterology & Motility* 29(4), e12969. doi: 10.1111/nmo.12969.

Igbinedion, S.O., Ansari, J., Vasikaran, A., Gavins, F.N., Jordan, P., Boktor, M., et al. (2017). Non-celiac gluten sensitivity: All wheat attack is not celiac. *World journal of gastroenterology* 23(40), 7201-7210. doi: 10.3748/wjg.v23.i40.7201.

Ivarsson, A., Persson, L., Nyström, L., Ascher, H., Cavell, B., Danielsson, L., et al. (2000). Epidemic of coeliac disease in Swedish children. *Acta Paediatrica* 89(2), 165-171. doi: 10.1111/j.1651-2227.2000.tb01210.x.

Junker, Y., Zeissig, S., Kim, S.-J., Barisani, D., Wieser, H., Leffler, D.A., et al. (2012). Wheat amylase trypsin inhibitors drive intestinal inflammation via activation of toll-like receptor 4. *Journal of experimental medicine* 209(13), 2395-2408. doi: 10.1084/jem.20102660.

Kalla, R., Shimamoto, K., Potter, R., Nielsen, P.S., Linnestad, C., and Olsen, O.-A. (1994). The promoter of the barley aleurone-specific gene encoding a putative 7 kDa lipid transfer protein confers aleurone cell-specific expression in transgenic rice. *The Plant Journal* 6(6), 849-860. doi: 10.1046/j.1365-313X.1994.6060849.x.

Karell, K., Louka, A.S., Moodie, S.J., Ascher, H., Clot, F., Greco, L., et al. (2003). Hla types in celiac disease patients not carrying the DQA1*05-DQB1*02 (DQ2) heterodimer: results from the european genetics cluster on celiac disease. *Human Immunology* 64(4), 469-477. doi: 10.1016/S0198-8859(03)00027-2.

Kasarda, D.D. (2013). Can an Increase in Celiac Disease Be Attributed to an Increase in the Gluten Content of Wheat as a Consequence of Wheat Breeding? *Journal of Agricultural and Food Chemistry* 61(6), 1155-1159. doi: 10.1021/jf305122s.

Keet, C.A., Matsui, E.C., Dhillon, G., Lenehan, P., Paterakis, M., and Wood, R.A. (2009). The natural history of wheat allergy. *Annals of Allergy, Asthma and Immunology* 102(5), 410-415. doi: 10.1016/S1081-1206(10)60513-3.

Khalid, A.N., and McMains, K.C. (2016). Gluten sensitivity: fact or fashion statement? Current Opinion in Otolaryngology & Head and Neck Surgery 24(3), 238-240. doi: 10.1097/moo.0000000000000263.

Kim, Y.G., Cha, J., and Chandrasegaran, S. (1996). Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proceedings of the National Academy of Sciences* 93(3), 1156-1160. doi: 10.1073/pnas.93.3.1156.

Kooy-Winkelaar, Y., van Lummel, M., Moustakas, A.K., Schweizer, J., Mearin, M.L., Mulder, C.J., et al. (2011). Gluten-Specific T Cells Cross-React between HLA-DQ8 and the HLA-DQ2α/DQ8β Transdimer. *The Journal of Immunology* 187(10), 5123. doi: 10.4049/jimmunol.1101179.

Larré, C., Lupi, R., Gombaud, G., Brossard, C., Branlard, G., Moneret-Vautrin, D.A., et al. (2011). Assessment of allergenicity of diploid and hexaploid wheat genotypes: Identification of allergens in the albumin/globulin fraction. *Journal of Proteomics* 74(8), 1279-1289. doi: 10.1016/j.jprot.2011.03.014.

Larretxi, I., Simon, E., Benjumea, L., Miranda, J., Bustamante, M.A., Lasa, A., et al. (2019). Gluten-free-rendered products contribute to imbalanced diets in children and adolescents with celiac disease. *European Journal of Nutrition* 58(2), 775-783. doi: 10.1007/s00394-018-1685-2.

Lebwohl, B., Ludvigsson, J.F., and Green, P.H.R. (2015). Celiac disease and non-celiac gluten sensitivity. *BMJ : British Medical Journal* 351, h4347. doi: 10.1136/bmj.h4347.

Lebwohl, B., Sanders, D.S., and Green, P.H.R. (2018). Coeliac disease. *The Lancet* 391(10115), 70-81. doi: 10.1016/S0140-6736(17)31796-8.

Li, T., Huang, S., Jiang, W.Z., Wright, D., Spalding, M.H., Weeks, D.P., et al. (2010). TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. *Nucleic Acids Research* 39(1), 359-372. doi: 10.1093/nar/gkq704.

LOHI, S., MUSTALAHTI, K., KAUKINEN, K., LAURILA, K., COLLIN, P., RISSANEN, H., et al. (2007). Increasing prevalence of coeliac disease over time. *Alimentary Pharmacology and Therapeutics* 26(9), 1217-1225. doi: 10.1111/j.1365-2036.2007.03502.x.

Losurdo, G., Principi, M., Iannone, A., Amoruso, A., Ierardi, E., Di Leo, A., et al. (2018). Extra-intestinal manifestations of non-celiac gluten sensitivity: An expanding paradigm. *World journal of gastroenterology* 24(14), 1521-1530. doi: 10.3748/wjg.v24.i14.1521.

Ludvigsson, J.F., Leffler, D.A., Bai, J.C., Biagi, F., Fasano, A., Green, P.H., et al. (2013). The Oslo definitions for coeliac disease and related terms. *Gut* 62(1), 43-52. doi: 10.1136/gutjnl-2011-301346.

Ma, X., Zhu, Q., Chen, Y., and Liu, Y.-G. (2016). CRISPR/Cas9 Platforms for Genome Editing in Plants: Developments and Applications. *Molecular Plant* 9(7), 961-974. doi: 10.1016/j.molp.2016.04.009.

Maiuri, L., Ciacci, C., Ricciardelli, I., Vacca, L., Raia, V., Auricchio, S., et al. (2003). Association between innate response to gliadin and activation of pathogenic T cells in coeliac disease. *The Lancet* 362(9377), 30-37. doi: 10.1016/S0140-6736(03)13803-2.

Maiuri, L., Picarelli, A., Boirivant, M., Coletta, S., Mazzilli, M.C., De Vincenzi, M., et al. (1996). Definition of the initial immunologic modifications upon *in vitro* gliadin challenge in the small intestine of celiac patients. *Gastroenterology* 110(5), 1368-1378. doi: 10.1053/gast.1996.v110.pm8613040.

Martín, A., Alvarez, J.B., Martín, L.M., Barro, F., and Ballesteros, J. (1999). The Development of Tritordeum: A Novel Cereal for Food Processing. *Journal of Cereal Science* 30(2), 85-95. doi: 10.1006/jcrs.1998.0235.

Matsuo, H., Morita, E., Tatham, A.S., Morimoto, K., Horikawa, T., Osuna, H., et al. (2004). Identification of the IgE-binding Epitope in ω-5 Gliadin, a Major Allergen in Wheat-dependent Exercise-induced Anaphylaxis. *Journal of Biological Chemistry* 279(13), 12135-12140. doi: 10.1074/jbc.M311340200.

Matsuo, H., Yokooji, T., and Taogoshi, T. (2015). Common food allergens and their IgE-binding epitopes. *Allergology International* 64(4), 332-343. doi: 10.1016/j.alit.2015.06.009.

Metakovsky, E., Wrigley, C., Bekes, F., Gupta, R., and Metakovskii, E. (1990). Gluten polypeptides as useful genetic markers of dough quality in Australian wheats. *Australian Journal of Agricultural Research* 41(2), 289-306. doi: 10.1071/AR9900289.

Mitea, C., Salentijn, E.M.J., van Veelen, P., Goryunova, S.V., van der Meer, I.M., van den Broeck, H.C., et al. (2010). A Universal Approach to Eliminate Antigenic Properties of Alpha-Gliadin Peptides in Celiac Disease. *PLOS ONE* 5(12), e15637. doi: 10.1371/journal.pone.0015637.

Molberg, Ø., Solheim flÆte, N., Jensen, T., Lundin, K.E.A., Arentz-Hansen, H., Anderson, O.D., et al. (2003). Intestinal T-cell responses to high-molecular-weight glutenins in celiac disease. *Gastroenterology* 125(2), 337-344. doi: 10.1016/S0016-5085(03)00890-4.

Molberg, Ø., Uhlen, A.K., Jensen, T., Flæte, N.S., Fleckenstein, B., Arentz–Hansen, H., et al. (2005). Mapping of gluten T-cell epitopes in the bread wheat ancestors: Implications for celiac disease. *Gastroenterology* 128(2), 393-401. doi: 10.1053/j.gastro.2004.11.003.

Morita, E., Matsuo, H., Mihara, S., Morimoto, K., Savage, A.W.J., and Tatham, A.S. (2003). Fast ω-gliadin is a major allergen in wheat-dependent exercise-induced anaphylaxis. *Journal of Dermatological Science* 33(2), 99-104. doi: 10.1016/S0923-1811(03)00156-7.

Mundy, J., Hejgaard, J., and Svendsen, I. (1984). Characterization of a bifunctional wheat inhibitor of endogenous α-amylase and subtilisin. *FEBS Letters* 167(2), 210-214. doi: 10.1016/0014-5793(84)80128-3.

Ozuna, C.V., and Barro, F. (2018). Characterization of gluten proteins and celiac disease-related immunogenic epitopes in the Triticeae: cereal domestication and breeding contributed to decrease the content of gliadins and gluten. *Molecular Breeding* 38(3), 22. doi: 10.1007/s11032-018-0779-0.

Ozuna, C.V., Iehisa, J.C.M., Giménez, M.J., Alvarez, J.B., Sousa, C., and Barro, F. (2015). Diversification of the celiac disease α-gliadin complex in wheat: a 33-mer peptide with six overlapping epitopes, evolved following polyploidization. *Plant Journal* 82(5), 794-805. doi: 10.1111/tpj.12851.

Palosuo, K., Varjonen, E., Nurkkala, J., Kalkkinen, N., Harvima, R., Reunala, T., et al. (2003). Transglutaminase-mediated cross-linking of a peptic fraction of ω-5 gliadin enhances IgE reactivity in wheat-dependent, exercise-induced anaphylaxis. *Journal of Allergy and Clinical Immunology* 111(6), 1386-1392. doi: 10.1067/mai.2003.1498.

Payne, P.I., Holt, L.M., Jarvis, M.G., and Jackson, E.A. (1985). Two-dimensional fractionation of the endosperm proteins of bread wheat (Triticum aestivum): Biochemical and genetic studies. *Cereal Chemistry* 62(5), 319-326.

Peters, S.L., Biesiekierski, J.R., Yelland, G.W., Muir, J.G., and Gibson, P.R. (2014). Randomised clinical trial: gluten may cause depression in subjects with non-coeliac gluten sensitivity – an exploratory clinical study. *Alimentary Pharmacology and Therapeutics* 39(10), 1104-1112. doi: 10.1111/apt.12730.

Petersen, G., Seberg, O., Yde, M., and Berthelsen, K. (2006). Phylogenetic relationships of Triticum and Aegilops and evidence for the origin of the A, B, and D genomes of common wheat (Triticum aestivum). *Molecular Phylogenetics and Evolution* 39(1), 70-82. doi: 10.1016/j.ympev.2006.01.023.

Picarelli, A., Borghini, R., Di Tola, M., Marino, M., Urciuoli, C., Isonne, C., et al. (2016). Intestinal, Systemic, and Oral Gluten-related Alterations in Patients With Nonceliac Gluten Sensitivity. *Journal of Clinical Gastroenterology* 50(10), 849-858. doi: 10.1097/MCG.0000000000000515.

Pinier, M., Fuhrmann, G., Galipeau, H.J., Rivard, N., Murray, J.A., David, C.S., et al. (2012). The Copolymer P(HEMA-co-SS) Binds Gluten and Reduces Immune Response in Gluten-Sensitized Mice and Human Tissues. *Gastroenterology* 142(2), 316-325.e312. doi: 10.1053/j.gastro.2011.10.038.

Pistón, F., Gil-Humanes, J., and Barro, F. (2013). Integration of promoters, inverted repeat sequences and proteomic data into a model for high silencing efficiency of coeliac disease related gliadins in bread wheat. *BMC plant biology* 13, 136-136. doi: 10.1186/1471-2229-13-136.

Pizzuti, D., Buda, A., D'Odorico, A., D'Incà, R., Chiarelli, S., Curioni, A., et al. (2006). Lack of intestinal mucosal toxicity of Triticum monococcum in celiac disease patients. *Scandinavian Journal of Gastroenterology* 41(11), 1305-1311. doi: 10.1080/00365520600699983.

Qiao, S.-W., Bergseng, E., Molberg, Ø., Jung, G., Fleckenstein, B., and Sollid, L.M. (2005). Refining the Rules of Gliadin T Cell Epitope Binding to the Disease-Associated DQ2 Molecule in Celiac Disease: Importance of Proline Spacing and Glutamine Deamidation. *The Journal of Immunology* 175(1), 254-261. doi: 10.4049/jimmunol.175.1.254.

Reese, I., Schäfer, C., Kleine-Tebbe, J., Ahrens, B., Bachmann, O., Ballmer-Weber, B., et al. (2018). Non-celiac gluten/wheat sensitivity (NCGS)—a currently undefined disorder without validated diagnostic criteria and of unknown prevalence. *Allergo Journal International* 27(5), 147-151. doi: 10.1007/s40629-018-0070-2.

Roberts, T.H., and Hejgaard, J. (2008). Serpins in plants and green algae. *Functional and Integrative Genomics* 8(1), 1-27. doi: 10.1007/s10142-007-0059-2.

Rosell, C.M., Barro, F., Sousa, C., and Mena, M.C. (2014). Cereals for developing gluten-free products and analytical tools for gluten detection. *Journal of Cereal Science* 59(3), 354-364. doi: 10.1016/j.jcs.2013.10.001.

Rubio-Tapia, A., and Murray, J.A. (2010). Classification and management of refractory coeliac disease. *Gut* 59(4), 547-557. doi: 10.1136/gut.2009.195131.

Rubio–Tapia, A., Kyle, R.A., Kaplan, E.L., Johnson, D.R., Page, W., Erdtmann, F., et al. (2009). Increased Prevalence and Mortality in Undiagnosed Celiac Disease. *Gastroenterology* 137(1), 88-93. doi: 10.1053/j.gastro.2009.03.059.

Ruiz-Carnicer, Á., Comino, I., Segura, V., Ozuna, C.V., Moreno, M.d.L., López-Casado, M.Á., et al. (2019). Celiac Immunogenic Potential of α-Gliadin Epitope Variants from Triticum and Aegilops Species. *Nutrients* 11(2), 220.

Sabelli, P.A., and Shewry, P.R. (1991). Characterization and organization of gene families at the Gli-1 loci of bread and durum wheats by restriction fragment analysis. *Theoretical and Applied Genetics* 83(2), 209-216. doi: 10.1007/BF00226253.

Salcedo, G., Quirce, S., and Diaz-Perales, A. (2011). Wheat allergens associated with Baker's asthma. *Journal of investigational allergology and clinical immunology* 21(2), 81-92; quiz 94.

Samanta, M.K., Dey, A., and Gayen, S. (2016). CRISPR/Cas9: an advanced tool for editing plant genomes. *Transgenic Research* 25(5), 561-573. doi: 10.1007/s11248-016-9953-5.

Sapone, A., Bai, J.C., Ciacci, C., Dolinsek, J., Green, P.H.R., Hadjivassiliou, M., et al. (2012). Spectrum of gluten-related disorders: consensus on new nomenclature and classification. *BMC Medicine* 10(1), 13. doi: 10.1186/1741-7015-10-13.

Sapone, A., Lammers, K.M., Mazzarella, G., Mikhailenko, I., Cartenì, M., Casolaro, V., et al. (2010). Differential Mucosal IL-17 Expression in Two Gliadin-Induced Disorders: Gluten Sensitivity and the Autoimmune Enteropathy Celiac Disease. *International Archives of Allergy and Immunology* 152(1), 75-80. doi: 10.1159/000260087.

Schuppan, D., and Zevallos, V. (2015). Wheat Amylase Trypsin Inhibitors as Nutritional Activators of Innate Immunity. *Digestive Diseases* 33(2), 260-263. doi: 10.1159/000371476.

Shan, L., Molberg, Ø., Parrot, I., Hausch, F., Filiz, F., Gray, G.M., et al. (2002). Structural Basis for Gluten Intolerance in Celiac Sprue. *Science* 297(5590), 2275. doi: 10.1126/science.1074129.

Shewry, P.R. (2009). Wheat. *Journal of Experimental Botany* 60(6), 1537-1553. doi: 10.1093/jxb/erp058.

Shewry, P.R., Halford, N.G., and Lafiandra, D. (2003). "Genetics of Wheat Gluten Proteins". *Advances in Genetics* 49, 111-184. doi: 10.1016/S0065-2660(03)01003-4

Shewry, P.R., Halford, N.G., and Tatham, A.S. (1992). High molecular weight subunits of wheat glutenin. *Journal of Cereal Science* 15(2), 105-120. doi: 10.1016/S0733-5210(09)80062-3.

Shewry, P.R., and Hey, S.J. (2015). The contribution of wheat to human diet and health. *Food and Energy Security* 4(3), 178-202. doi: 10.1002/fes3.64.

Shewry, P.R., Tatham, A.S., Forde, J., Kreis, M., and Miflin, B.J. (1986). The classification and nomenclature of wheat gluten proteins: A reassessment. *Journal of Cereal Science* 4(2), 97-106. doi: 10.1016/S0733-5210(86)80012-1.

Silverman, G.A., Bird, P.I., Carrell, R.W., Church, F.C., Coughlin, P.B., Gettins, P.G.W., et al. (2001). The Serpins Are an Expanding Superfamily of Structurally Similar but Functionally Diverse Proteins: Evolution, mechanism of inhibition, novel functions, and a revised nomenclature. *Journal of Biological Chemistry* 276(36), 33293-33296. doi: 10.1074/jbc.R100016200.

Singh, P., Arora, A., Strand, T.A., Leffler, D.A., Catassi, C., Green, P.H., et al. (2018). Global Prevalence of Celiac Disease: Systematic Review and Meta-analysis. *Clinical Gastroenterology and Hepatology* 16(6), 823-836.e822. doi: 10.1016/j.cgh.2017.06.037.

Sjöström, Lundin, Molberg, Körner, Mcadam, Anthonsen, et al. (1998). Identification of a Gliadin T-Cell Epitope in Coeliac Disease: General Importance of Gliadin Deamidation for Intestinal T-Cell Recognition. *Scandinavian Journal of Immunology* 48(2), 111-115. doi: 10.1046/j.1365-3083.1998.00397.x.

Skodje, G.I., Sarna, V.K., Minelle, I.H., Rolfsen, K.L., Muir, J.G., Gibson, P.R., et al. (2018). Fructan, Rather Than Gluten, Induces Symptoms in Patients With Self-Reported Non-Celiac Gluten Sensitivity. *Gastroenterology* 154(3), 529-539.e522. doi: 10.1053/j.gastro.2017.10.040.

Sollid, L.M. (2002). Coeliac disease: dissecting a complex inflammatory disorder. *Nature Reviews Immunology* 2(9), 647-655. doi: 10.1038/nri885.

Spaenij–Dekking, L., Kooy–Winkelaar, Y., van Veelen, P., Wouter Drijfhout, J., Jonker, H., van Soest, L., et al. (2005). Natural Variation in Toxicity of Wheat: Potential for Selection of Nontoxic Varieties for Celiac Disease Patients. *Gastroenterology* 129(3), 797-806. doi: 10.1053/j.gastro.2005.06.017.

Stepniak, D., Vader, L.W., Kooy, Y., van Veelen, P.A., Moustakas, A., Papandreou, N.A., et al. (2005). T-cell recognition of HLA-DQ2-bound gluten peptides can be influenced by an N-terminal proline at p-1. *Immunogenetics* 57(1), 8-15. doi: 10.1007/s00251-005-0780-8.

Šuligoj, T., Gregorini, A., Colomba, M., Ellis, H.J., and Ciclitira, P.J. (2013). Evaluation of the safety of ancient strains of wheat in coeliac disease reveals heterogeneous small intestinal T cell responses suggestive of coeliac toxicity. *Clinical Nutrition* 32(6), 1043-1049. doi: 10.1016/j.clnu.2013.02.003.

Sun, Y., Li, J., and Xia, L. (2016). Precise Genome Modification via Sequence-Specific Nucleases-Mediated Gene Targeting for Crop Improvement. *Frontiers in Plant Science* 7(1928). doi: 10.3389/fpls.2016.01928.

Tanner, G.J., Blundell, M.J., Colgrave, M.L., and Howitt, C.A. (2013). Quantification of Hordeins by ELISA: The Correct Standard Makes a Magnitude of Difference. *PLOS ONE* 8(2), e56456. doi: 10.1371/journal.pone.0056456.

Tatham, A.S., and Shewry, P.R. (2008). Allergens to wheat and related cereals. *Clinical and Experimental Allergy* 38(11), 1712-1726. doi: 10.1111/j.1365-2222.2008.03101.x.

Tatham, A.S., and Shewry, P.R. (2012). The S-poor prolamins of wheat, barley and rye: Revisited. *Journal of Cereal Science* 55(2), 79-99. doi: 10.1016/j.jcs.2011.10.013.

Tjon, J.M., van Bergen, J., and Koning, F. (2010). Celiac disease: how complicated can it get? *Immunogenetics* 62(10), 641-651. doi: 10.1007/s00251-010-0465-9.

Tollefsen, S., Arentz-Hansen, H., Fleckenstein, B., Molberg, O., Ráki, M., Kwok, W.W., et al. (2006). HLA-DQ2 and -DQ8 signatures of gluten T cell epitopes in celiac disease. *The Journal of clinical investigation* 116(8), 2226-2236. doi: 10.1172/JCI27620.

Tovoli, F., Masi, C., Guidetti, E., Negrini, G., Paterini, P., and Bolondi, L. (2015). Clinical and diagnostic aspects of gluten related disorders. *World journal of clinical cases* 3(3), 275-284. doi: 10.12998/wjcc.v3.i3.275.

Tye-Din, J.A., Stewart, J.A., Dromey, J.A., Beissbarth, T., van Heel, D.A., Tatham, A., et al. (2010b). Comprehensive, Quantitative Mapping of T Cell Epitopes in Gluten in Celiac Disease. *Science Translational Medicine* 2(41), 41ra51-41ra51. doi: 10.1126/scitranslmed.3001012.

Uhde, M., Ajamian, M., Caio, G., De Giorgio, R., Indart, A., Green, P.H., et al. (2016). Intestinal cell damage and systemic immune activation in individuals reporting sensitivity to wheat in the absence of coeliac disease. *Gut* 65(12), 1930-1937. doi: 10.1136/gutjnl-2016-311964.

Vader, L.W., Stepniak, D.T., Bunnik, E.M., Kooy, Y.M.C., De Haan, W., Drijfhout, J.W., et al. (2003). Characterization of cereal toxicity for celiac disease patients based on protein homology in grain. *Gastroenterology* 125(4), 1105-1113. doi: 10.1016/S0016-5085(03)01204-6.

Vader, W., Kooy, Y., van Veelen, P., de Ru, A., Harris, D., Benckhuijsen, W., et al. (2002). The gluten response in children with celiac disease is directed toward multiple gliadin and glutenin peptides. *Gastroenterology* 122(7), 1729-1737. doi: 10.1053/gast.2002.33606.

van de Wal, Y., Kooy, Y.M., van Veelen, P.A., Peña, S.A., Mearin, L.M., Molberg, O., et al. (1998). Small intestinal T cells of celiac disease patients recognize a natural pepsin fragment of gliadin. *Proceedings of the National Academy of Sciences of the United States of America* 95(17), 10050-10054. doi: 10.1073/pnas.95.17.10050.

van de Wal, Y., Kooy, Y.M.C., van Veelen, P., Vader, W., August, S.A., Drijfhout, J.W., et al. (1999). Glutenin is involved in the gluten-driven mucosal T cell response. *European Journal of Immunology* 29(10), 3133-3139. doi: 10.1002/(sici)1521-4141(199910)29:10<3133::aid-immu3133>3.0.co;2-g.

van den Broeck, H., Hongbing, C., Lacaze, X., Dusautoir, J.-C., Gilissen, L., Smulders, M., et al. (2010). In search of tetraploid wheat accessions reduced in celiac disease-related gluten epitopes. *Molecular BioSystems* 6(11), 2206-2213. doi: 10.1039/C0MB00046A.

van den Broeck, H.C., van Herpen, T.W.J.M., Schuit, C., Salentijn, E.M.J., Dekking, L., Bosch, D., et al. (2009). Removing celiac disease-related gluten proteins from bread wheat while retaining technological properties: a study with Chinese Spring deletion lines. *BMC Plant Biology* 9(1), 41. doi: 10.1186/1471-2229-9-41.

van Herpen, T.W.J.M., Goryunova, S.V., van der Schoot, J., Mitreva, M., Salentijn, E., Vorst, O., et al. (2006). Alpha-gliadin genes from the A, B, and D genomes of wheat contain different sets of celiac disease epitopes. *BMC Genomics* 7(1), 1. doi: 10.1186/1471-2164-7-1.

Vaquero, L., Alvarez-Cuenllas, B., Rodríguez-Martín, L., Aparicio, M., Jorquera, F., Olcoz, J.L., et al. (2015). Revisión de las patologías relacionadas con la ingesta de gluten. *Nutrición Hospitalaria* 31, 2359-2371.

Vaquero, L., Comino, I., Vivas, S., Rodríguez-Martín, L., Giménez, M.J., Pastor, J., et al. (2018). Tritordeum: a novel cereal for food processing with good acceptability and significant reduction in gluten immunogenic peptides in comparison with wheat. *Journal of the Science of Food and Agriculture* 98(6), 2201-2209. doi: 10.1002/jsfa.8705.

Vazquez-Roque, M., and Oxentenko, A.S. (2015). Nonceliac Gluten Sensitivity. *Mayo Clinic Proceedings* 90(9), 1272-1277. doi: 10.1016/j.mayocp.2015.07.009.

Ventura, A., Magazzù, G., and Greco, L. (1999). Duration of exposure to gluten and risk for autoimmune disorders in patients with celiac disease. *Gastroenterology* 117(2), 297-303. doi: 10.1053/gast.1999.0029900297.

Vici, G., Belli, L., Biondi, M., and Polzonetti, V. (2016). Gluten free diet and nutrient deficiencies: A review. *Clinical Nutrition* 35(6), 1236-1241. doi: 10.1016/j.clnu.2016.05.002.

Vojdani, A., and Vojdani, E. (2017). Gluten and non-gluten proteins of wheat as target antigens in autism, Crohn's and celiac disease. *Journal of Cereal Science* 75, 252-260. doi: 10.1016/j.jcs.2017.04.010.

Volta, U., Pinto-Sanchez, M.I., Boschetti, E., Caio, G., De Giorgio, R., and Verdu, E.F. (2016). Dietary Triggers in Irritable Bowel Syndrome: Is There a Role for Gluten? *Journal of neurogastroenterology and motility* 22(4), 547-557. doi: 10.5056/jnm16069.

Wen, S., Wen, N., Pang, J., Langen, G., Brew-Appiah, R.A.T., Mejias, J.H., et al. (2012). Structural genes of wheat and barley 5-methylcytosine DNA glycosylases and their potential applications for human health. *Proceedings of the National Academy of Sciences* 109(50), 20543-20548. doi: 10.1073/pnas.1217927109.

Wieser, H. (2007). Chemistry of gluten proteins. *Food Microbiology* 24(2), 115-119. doi: 10.1016/j.fm.2006.07.004.

Wieser, H., and Kieffer, R. (2001). Correlations of the Amount of Gluten Protein Types to the Technological Properties of Wheat Flours Determined on a Micro-scale. *Journal of Cereal Science* 34(1), 19-27. doi: 10.1006/jcrs.2000.0385.

Xie, K., Minkenberg, B., and Yang, Y. (2015). Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proceedings of the National Academy of Sciences* 112(11), 3570-3575. doi: 10.1073/pnas.1420294112.

Xie, Z., Wang, C., Wang, K., Wang, S., Li, X., Zhang, Z., et al. (2010). Molecular characterization of the celiac disease epitope domains in α-gliadin genes in Aegilops tauschii and hexaploid wheats (Triticum aestivum L.). *Theoretical and Applied Genetics* 121(7), 1239-1251. doi: 10.1007/s00122-010-1384-8.

Xing, H.-L., Dong, L., Wang, Z.-P., Zhang, H.-Y., Han, C.-Y., Liu, B., et al. (2014). A CRISPR/Cas9 toolkit for multiplex genome editing in plants. *BMC Plant Biology* 14(1), 327. doi: 10.1186/s12870-014-0327-y.

Zanini, B., Baschè, R., Ferraresi, A., Ricci, C., Lanzarotto, F., Marullo, M., et al. (2015). Randomised clinical study: gluten challenge induces symptom recurrence in only a minority of patients who meet clinical criteria for non-coeliac gluten sensitivity. *Alimentary Pharmacology and Therapeutics* 42(8), 968-976. doi: 10.1111/apt.13372.

CHAPTER 2

Tritordeum Breads are well Tolerated with Preference over Gluten-Free Breads in Non-Celiac Gluten-Sensitive (NCGS) Patients and its Consumption Induce Changes in Beneficial Short-Chain Fatty Acids-Producing Gut Bacteria.

Tritordeum Breads are well Tolerated with Preference over Gluten-Free Breads in Non-Celiac Gluten-Sensitive (NCGS) Patients and its Consumption Induce Changes in Beneficial Short-Chain Fatty Acids-Producing Gut Bacteria.

Abstract

Cereals, in particular wheat, are an important source for carbohydrate, proteins, fiber as well as other compounds. However, the ingestion of wheat is also related with several gut disorders, among which non-celiac gluten sensitivity (NCGS) may affect up to 13% of population. The specific components responsible for NCGS have still to be confirmed and may include gluten and other wheat, barley and rye compounds. Tritordeum is a new cereal derived from crossing durum wheat with a wild barley species, which has important health-related traits differing from wheat in its gluten content. In the present work we examined the response of patients diagnosed as NCGS to tritordeum breads, to test if they could tolerate this new cereal in their diet. The trial was performed in two phases ("Basal" and "Tritordeum"), each of which lasted seven days. Gastrointestinal symptoms as well as tritordeum acceptability were recorded, Gluten Immunogenic Peptides excretion was evaluated and the composition and structure of the intestinal microbiota was analyzed by sequencing of the 16S rRNA at the end of each phase. Gastrointestinal symptoms of the subjects showed no significant change between the Basal phase (gluten free bread) and the Tritordeum phase. As expected, there was a low excretion of GIP by subjects during the Basal phase. The trial subjects rated tritordeum bread higher than the Basal (gluten-free) bread; in particular, texture and taste were scored highly in tritordeum breads. Analysis of the bacterial gut microbiota indicated that its consumption does not alter the global structure and composition of the intestinal microbiota, remaining stable after dietary intervention in these patients. However, a significantly higher abundance of the butyrate-producing *Firmicutes Faecalibacterium* genus and a significantly lower abundance of the *Ruminococcaceae UCG—013* genus, was induced in the tritordeum phase compared with the Basal phase. The increase in the relative abundance of *Faecalibacterium* genus is important as it has been shown to be one of the main producing genera of butyrate, a short chain fatty acid known to have a positive effect on human health, especially maintaining intestinal homeostasis and immune tolerance. The results suggest that tritordeum may be tolerated by at least a sub-set of NCGS sufferers who do not require strict exclusion of gluten from their diet.

Introduction

The cereals, and in particular wheat, have a central role in human diet and are important sources of complex carbohydrate, protein and dietary fibre as well as minerals, vitamins and phytoactive compounds (Shewry and Hey, 2015). They have highly valued organoleptic properties and are functionally versatile allowing them to be use in a great number of food applications. Although in general the cereals are very safe foodstuffs consumed routinely by billions of people, they are implicated in some food allergies and dietary intolerances, mainly related to their protein composition and in particular to gluten proteins present in cereals such as wheat, barley and rye. Gluten-related disorders include wheat allergy (WA) an allergic reaction associated with a spectrum of different proteins, including gliadins and amylase trypsin inhibitors (ATIs) (Tatham and Shewry, 2008); celiac disease (CD), an immune-mediated enteropathy induced by exposure to dietary gluten and non-celiac gluten sensitivity (NCGS) a disease related to the ingestion of gluten-containing foods in persons who do not suffer from WA or CD (Lebwohl et al., 2015; Dale et al., 2019). The specific components responsible for NCGS have still to be confirmed and may include gluten proteins, ATIs, fermentable mono- and polysaccharides and polyols (FODMAPS) or other factors (Dale et al., 2019), but with an estimated frequency of 6 – 13% the number of persons affected by NCGS is significantly higher than WA (around 1-2%) or CD (around 1%) (Sapone et al., 2012; Lebwohl et al., 2015; Aziz et al., 2016; Dale et al., 2019). Although there is evidence that not all NCGS sufferers need to exclude gluten completely from their diet as they may have different levels of tolerance (Roncoroni et al., 2019), many follow a gluten-free diet (GFD), although this may have negative nutritional implications as GFDs containing more fat and less carbohydrate than recommended and tend to be poorer in several important dietary components such as folic acid, calcium, iron and vitamin B6, B12 and D (Gobbetti et al., 2018; Larretxi et al., 2019). In addition a strict GFD may have negative effects on beneficial gut microbiota (De Palma et al., 2009; Lorenzo Pisarello et al., 2015; Haro et al., 2018). In fact, a microbial dysbiosis in NCGS causes gut inflammation, abdominal pain, metabolic problems, neuroinflammation, gut-brain axis dysfunction, intestinal barrier defects and inflammatory responses to gluten, among others (Daulatzai, 2015; Volta et al., 2017). In addition, Uhde et al. (Uhde et al., 2016) demonstrated the existence of an increased intestinal permeability in subjects with non-coeliac/non-allergic gluten/wheat sensitivity (NCG/WS) indicating a possible key pathogenic role of the intestinal microbiota in these patients. In this sense, it is known that bacterial species mainly belonging to *Firmicutes* phylum, perform important immunological, structural and metabolic functions in the host, among which, it should be noted its involvement in the preservation of the integrity of the intestinal barrier through the production of short-chain fatty acids (SCFA) (Brenchley and Douek, 2012).

Among the SCFA, butyrate seems to have a greater contribution in this function, because it constitutes the energy source of the enterocytes, keep the tight junctions, stimulates mucin production and therefore favours the maintenance of the intestinal mucus layer avoiding

intestinal permeability, and finally it induces the production of antimicrobial peptides (Hatayama et al., 2007; Canani et al., 2011). Taken together, all these functions of butyrate ultimately result in intestinal homeostasis and immune tolerance (Leccioli et al., 2017; Yan and Ajuwon, 2017). In this context, it is of interest to explore whether NCGS sufferers may safely consume cereal foods modified by techniques such as sourdough fermentation which can reduce levels of gluten proteins and FODMAPS (Calasso et al., 2018; Muir et al., 2019) or made using alternative gluten sources which may be better tolerated as they have lower levels of symptom-inducing components (Vaquero et al., 2018). In a previous study, the present authors showed that tritordeum, a new cereal crop species derived from crossing durum wheat with a wild barley species (Cubero et al., 1986; Martín et al., 1999) differs from bread wheat in its gluten composition, in particular with respect to gliadin composition, showing significantly lower levels of gluten immunogenic peptides (GIP) implicated in CD and possibly also in NCGS. In addition, it was seen that healthy subjects who consumed a tritordeum diet during one week showed a significant reduction of the excretion of GIP in faeces by comparison with a bread wheat diet (Vaquero et al., 2018). The present study follows on from the previous work, with the aim of examining the response of patients diagnosed as NCGS to tritordeum breads, to test the hypothesis that the reduced GIP levels in tritordeum may allow some NCGS patients to tolerate it in their diet. Additionally, the study analyses the effect of its consumption on the intestinal microbiota of these patients, due to the recent idea, that a dysbiosis of the intestinal microbiota could contribute to the aetiology and pathogenesis of NCGS (Daulatzai, 2015).

Materials and Methods

Subjects

The subjects for the trial were twelve adult volunteers (9 / 12 female) of age 31-57 years, diagnosed as suffering from NCGS that fulfilled the Salerno criteria (Catassi et al., 2015). All subjects had been diagnosed NCGS at least six months before the start of the trial and during the six months previous to the trial they had been following a gluten-free diet (GFD) without interruption. All subjects showed negative serology for celiac disease (tissue-transglutaminase IgA antibodies) and the duodenal biopsy results presented normal duodenal villi architecture. HLA-DQ2+ was present in five of the subjects, whilst the remaining seven showed a different HLA-DQ2 or DQ8 haplotype. The subjects did not suffer from any other chronic disease and were not taking long-term medication. Of the twelve subjects who started the trial, two withdrew voluntarily during the first days of the trial (during the Basal, GFD phase), as they were not able to attend to deliver stool samples (the two were female with HLA-DQ2 or DQ8 not present). Their withdrawal was not associated with any change in gastro-intestinal symptoms or any other physiological reason. The remaining ten subjects followed the two phases of the trial, completed the clinical and sensory questionnaires and provided the required stool samples.

All study participants provided informed consent, and the study design was approved by the Ethics Review Board of the Hospital of León, Spain (Approval Number 1626).

Study design

The trial was performed in two phases ("Basal" and "Tritordeum"), each of which had duration of seven days. At the beginning of the trial, subjects were given a hand-out explaining the background / objectives of the trial and giving instructions for the handling and consumption of the bread samples and the collection and delivery of stool samples.

Basal Phase: consumption of the GFD normally eaten by the subject including the consumption of the gluten-free bread habitually consumed by each individual. At the end of the phase, Clinical Questionnaires based on Gastrointestinal Symptom Rating Scale (GSRS) and Sensory Questionnaires on palatability and acceptability of the bread consumed (Vaquero et al., 2018) were completed and stool samples (Basal Phase samples) were collected.

Tritordeum Phase: Continuation of the Basal GFD, but with the substitution of the basal (gluten-free) bread by tritordeum bread during seven days. Subjects were instructed to consume a minimum of 100 g and a maximum of 150 g of tritordeum bread daily, by the consumption of four slices of bread, giving a daily intake of ca. 5 - 7.5 g of gluten, based on the gluten content of the same tritordeum breads previously reported (Vaquero et al., 2018). At the end of this phase, the Clinical and Sensory Questionnaires were completed and stool samples (Tritordeum Phase samples) were collected.

Preparation and provision of test breads

The ingredients for the tritordeum breads provided to subjects in the Tritordeum Phase were: flour, water, sourdough, salt, and baker's yeast. Breads were made using tritordeum sourdough, prepared 24 h before bread-making, by mixing tritordeum flour, water and sourdough starter culture (composed of flour, water, lactic acid bacteria and sourdough yeasts). The mixture was allowed to stand at room temperature for 3 h before storage at 5 ºC until use. For bread-making, the main dough got 150g Kg of tritordeum sourdough, and mixed with additional tritordeum flour, water, salt and baker yeast to form the final dough. This dough rested for 90 min at room temperature (22-24ºC). After resting bulk dough was divided into pieces, which were placed into metal loaf molds and allowed to ferment until they achieved 2/3 of the mold volume (approximately 90 min). Loaves were baked at 210 ºC - for 45 min. The acidity values of the tritordeum sourdough and tritordeum bread are the following; Tritordeum Sourdough: TTA 13, pH 4.05, Tritordeum Bread: pH 4.9

After cooling, loaves were cut into slices and frozen in portions. The bread was stored at the Hospital of León, where the subjects received the bread for the Tritordeum

Phase. The breads were supplied frozen to test subjects for defrosting immediately before consumption.

Evaluation of gastrointestinal symptoms

Gastrointestinal symptoms were reported by trial subjects using the GSRS questionnaire, a validated, self-administered questionnaire that includes 15 questions, which assess gastrointestinal symptoms using a 7-point Likert scale in five domains: Indigestion, Diarrhea, Constipation, Abdominal Pain and Reflux. The severity of symptoms reported in the GSRS increases with increasing score.

Evaluation of acceptability of breads

The Sensory Questionnaire required the subjects to score the bread consumed during the Basal and Tritordeum phases under five attributes: Appearance; Aroma; Crumb Texture; Taste; and Overall Acceptability, using a numerical scale of 1–9, each number corresponding to a description of the level of liking / dislike of the bread.

Collection of stool samples

Subjects were instructed to collect a 2-4 g stool sample into a sealed container after recording their food intake for seven days. Specimens were delivered within the first two hours after deposition and were stored at -80 ºC until processing. All samples were identified and labelled with a randomized numeric code.

Quantification of GIP in stool samples

In the study a total of 20 stool samples were analyzed from 10 subjects. The concentration of GIP in stools was measured by sandwich ELISA (iVYDAL In Vitro Diagnostics® iVYLISA GIP-S kit, Biomedal S.L., Seville, Spain) following the manufacturer's guidelines (Vaquero et al., 2018). Briefly, stool samples were mixed with 9 ml Universal Gluten Extraction Solution (UGES) per gram of stool then incubated at 50ºC for 60 min with gentle agitation to release the GIP from the stool matrix. After extraction, samples were diluted 1:10 with dilution solution and ELISA was performed using the provided G12 coated microtiter plate, standards and positive and negative controls. Each sample was run in duplicate and at least two different aliquots of each sample were tested on different days.

Sequencing and Bioinformatics Analysis of the gut microbiota

For each DNA of 36 fecal samples (9 patients per two phase of study and two DNA repetitions for each phase), the V1-V2 hypervariable regions of the bacterial 16S rRNA gene was amplified by polymerase chain reaction (PCR) using the universal bacterial primers 8F

(AGAGTTTGATCMTGGCTCAG) and 357R (CTGCTGCCTYCCGTA). In addition, 8nt index and Illumina adapter sequences were added following manufacturer's instructions. The PCR experimental conditions and single-end sequencing procedures were the same as in a previous study (Santos-Marcos et al., 2018). The raw NGS data from this study can be found at GenBank database under accession number PRJNA577543.

The Illumina Miseq Fastq reads obtained were analysed using the Quantitative Insights into Microbial Ecology 2 (QIIME 2) pipeline (version v2019.7) (Bolyen et al., 2019) with default parameters unless otherwise noted. Reads were processed by the DADA2 program using the qiime dada2 denoise-single script (Callahan et al., 2016) which denoises single-end sequences, dereplicates them, and filters chimeras. Open reference operational taxonomic unit (OTU) picking was performed using VSEARCH (Rideout et al., 2014) and the SILVA v132 reference databases at 97% identity, which provides a feature table containing the frequencies of each OTU or taxon per sample (Quast et al., 2013).

Rarefaction curves of alpha-diversity indexes (Faith_pd, Shannon, Observed OTUS and Good's coverage) and beta diversity (Unweighted and Weighted UniFrac distances) were calculated at an even sampling depth of 10,044 sequences per sample and used to assess differences in microbial diversity between the two phases of the study (https://github.com/qiime2/q2-diversity). Finally, taxonomic and compositional analyses were conducted by using the plugins feature-classifier classify-consensus-vsearch (Rognes et al., 2016) and taxa barplot (https://github.com/qiime2/q2-taxa).

Statistical analysis

The ANOVA test was performed to evaluate differences between various samples, followed by the two-tailed Dunnett's post hoc test for multiple mean comparisons. The Student's t test, Wilcoxon test and Friendman test were employed to analyse differences in the same subject among the different phases. The Bonferroni test was applied in the post hoc analysis and p values lower than 0.05 were considered statistically significant.

A non-supervised multivariate hierarchical clustering analysis, using Euclidean distance and the Ward clustering algorithm, and a supervised principal least square-discriminant analysis (PLS-DA) of all bacterial taxa from the different periods of dietary intervention were performed using MetaboAnalyst 4.01 (Chong et al., 2019).

Differences in alpha diversity indexes were estimated using the non-parametric Kruskal Wallis test and differences for the beta diversity unweighted and weighted UNIFRAC distances were estimated using the PERMANOVA analysis. Finally, the differences in the relative abundance of bacterial taxa between the two phases were tested using the non-parametric Mann-Whitney U test with SPSS Statistics for Windows Version 25.0 (IBM Corp., Armonk, NY). For that only taxa that were present in at least 80% of the samples per each phase were used.

Results

Gastro-intestinal symptoms

The results from the completion of the GSRS questionnaires, evaluating the parameters Indigestion, Diarrhea, Constipation, Abdominal Pain and Reflux, at the end of the Basal and Tritordeum phases indicated that the gastrointestinal symptoms of the subjects showed no significant change between the Basal phase (gluten free bread) and the Tritordeum phase (Figure 1). The global mean value (data for all subjects) during the Basal phase was 36.5 *vs* 34.1 during the Tritordeum phase (with higher values equating to lower gastrointestinal wellbeing) but these scores were not significantly different (P = 0.414). Examination of the questionnaire results for individual subjects showed that at the individual level there were no significant differences in gastrointestinal health scores (Figure 1).

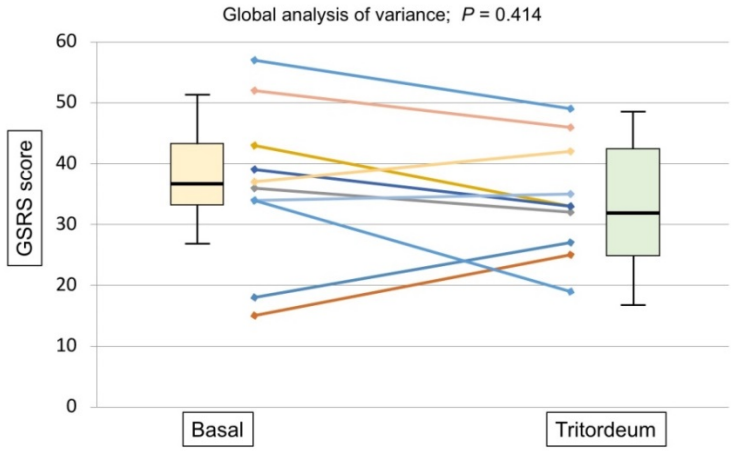


Figure 1. Comparison of the Symptom Rating Scale (GSRS) questionnaire for all patients (boxes) and results for individual subjects (each line represent one individual subject).

Fecal GIP content

The analysis of GIP in stool samples showed that, as expected, there was a low excretion of GIP (mean value 0.6 (0 – 2.9) µg GIP/g feces) by subjects during the Basal phase, although two subjects showed outstanding excretion levels (Figure 2). Higher levels of GIPs were detected in stool samples from the Tritordeum phase, confirming their consumption of gluten. The mean level of GIP detected in stool samples in the Tritordeum phase, was 1.5 µg GIP/g feces, with concentration ranged from 0.5 to 3.2, although this difference was not statistically significant (P = 0.114). The results obtained in the Tritordeum phase contrasted with those reported previously for subjects who included

standard wheat breads containing gluten in their diet, GIP concentration ranged from 0.30 to 75.68 µg/g feces (Vaquero et al., 2018).

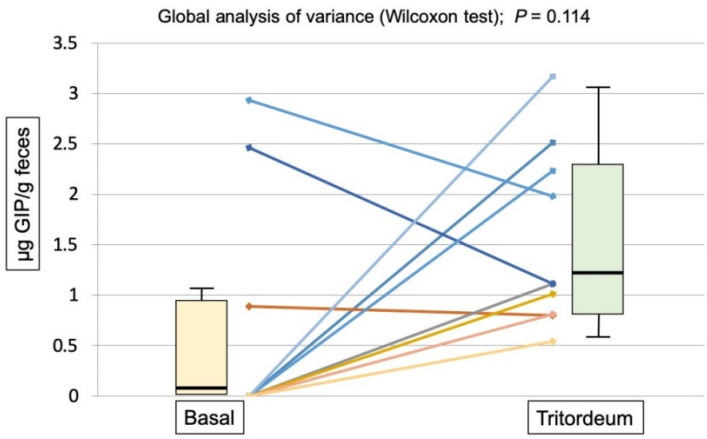


Figure 2. Comparison of the fecal Gluten Immunogenic Peptides (GIP) content after consumption of gluten-free bread (Basal) and Tritordeum breads.

Acceptability

Based on the Sensory Questionnaire, the trial subjects rated tritordeum bread higher than the basal (gluten-free) bread they habitually consumed for each of the five attributes scored (Appearance, Aroma, Texture, Taste and Overall Acceptability) and with a significant difference (P = < 0.05) seen with each attribute (Figure 3). In particular, Texture and Taste were scored highly in tritordeum breads in comparison with gluten-free breads (Tritordeum mean 7.3 *vs* Basal mean 3.6 and Tritordeum mean 7.7 *vs* Basal mean 4.6, respectively). The mean Overall Acceptability score for tritordeum bread was 7.2 *vs* 5.1 for basal breads (Figure 3).

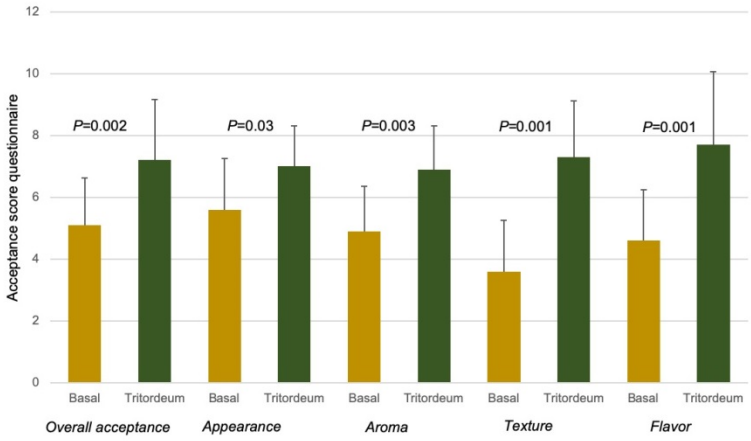


Figure 3. Differences in descriptive sensory analysis of the two phases of the dietary intervention. Values of the sensory questionnaire for Overall acceptance, Appearance, Aroma, Texture, and Flavor of the sensory questionnaire for the gluten-free bread (Basal) and Tritordeum bread. Data are the means of ten replications +/- standard deviation.

Diversity analysis of the gut microbiota

After processing of reads a total of 361,584 sequences were retained and a sample was lost. The sampling depth selected for rarefication was 10,044. Satisfactory Good´s coverage of the diversity was obtained for all samples since the mean value was > 99.9% for al samples. We did not find any significant differences in bacterial diversity between the Basal phase and the Tritordeum phase, with any of the alpha diversity indexes used (Faith_pd, $P = 0.480$; Shannon, $P = 0.635$, and observed OTUS, $P = 0.701$) (Figure 4A). In the same way, we did not find any significant differences in the beta diversity between both phases both for Unweighted ($P = 0.429$) and Weighted ($P = 0.461$) Unifrac distances. Similarly, no significant differences in both Unweighted and Weighted Unifrac distances were found between both phases for each patient when analysed independently ($P > 0.091$; Figure 4B).

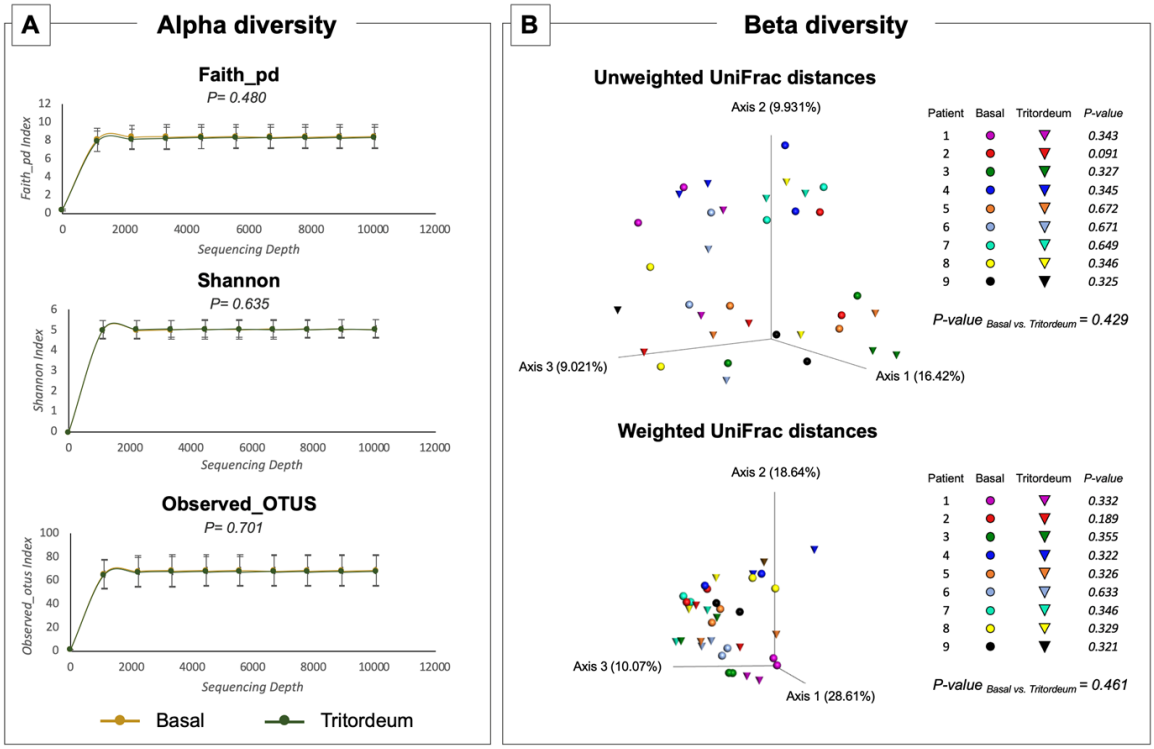


Figure 4. Alpha and Beta diversity analysis results. A) Alpha rarefaction curves of Faith_pd, Shannon and Observed otus alpha diversity indexes for the basal and Tritordeum phase at a depth of 10044 sequences per sample. *P*-value for each alpha diversity index was estimated using Kruskal-Wallis test. B) Principal Coordinate Analysis (PCoA) of the unweighted and weighted UniFrac beta diversity distances between the basal and Tritordeum phase at a depth of 10044 sequences per sample. P-value was estimated using a Permanova test comparing the basal and Tritordeum phases globally or for each patient independently.

The hierarchical clustering analysis showed a trend to group the two dietary phases of each patient in pairs, indicating the maintenance of the structure and composition of the intestinal microbiota of each patient after consumption with tritordeum (Supplementary Figure 2). PLS-DA of all bacterial taxa showed a good separation of both dietary phases ($P <$

0.01; Figure 5A) and ranked 15 bacterial taxa as the most importance in projection (VIP scores > 1.5 at $P < 0.05$; Figure 5B), of which eight bacterial taxa are SCFA bacteria.

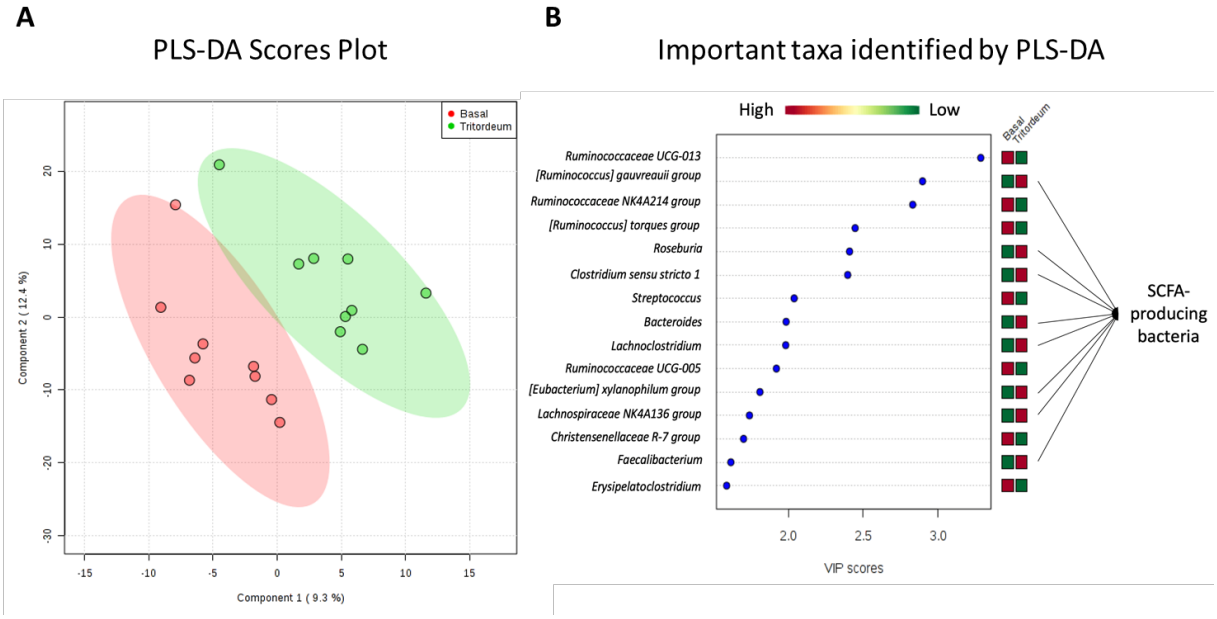


Figure 5. Partial Least Squares - Discriminant Analysis (PLS-DA). A) Partial least square-discriminant analysis (PLS-DA) 2D scores plot of the bacterial taxa of each phase of study. Red: basal phase; Green: Tritordeum phase. The model was established using two principal components; explained variance is in parentheses. B) Loading importance of bacterial taxa in the first PLS-DA component. Colored boxes (red: basal phase; green: Tritordeum phase) indicate relative concentrations of the corresponding bacterial taxa in each phase. VIP: Variable Importance in Projection. SCFA; Short-short-chain fatty acids.

Relative Abundance, and Changes of the Gut Microbiota

In total, 3 phyla, 5 classes, 5 orders, 9 families and 17 genera, were identified with a mean relative abundance equal to or greater than 0.1% in both phases and with a presence equal to or greater than 65% in all samples for each phase.

We did not find significant ($P > 0.05$) differences in the relative abundance of any taxa at phylum, class, order and family level between Basal and Tritordeum phases. On average, the taxa with the highest abundance both at the Basal phase and the Tritordeum phase were: *Firmicutes* 89.2% - 90.1% and A*ctinobacteria* 3.7% - 3.2% at the phylum level, *Clostridia* 69.9% - 72.2%, and *Erysipelotridia* 12.1% - 10.7% at class level, and *Clostridiales* 69.9% - 72.2% and *Eryselotrichales* 12.1% - 10.7% at order level and *Lachnospiraceae* 36.3% - 38.7% and *Ruminococcaceae* 26.1% - 23.9% at family level (Supplementary Figure 1).

Finally, at genus level we found significant differences in the relative abundance of two bacterial genera. We have observed a higher relative abundance in *Ruminococcaceae UCG-013* ($P = 0.024$) and a lower relative abundance in *Faecalibacterium* ($P = 0.011$) genera in

Basal phase as compared with tritordeum phase (Figure 5). These results are in agreement with those observed in the PLS-DA analysis (Figure 6).

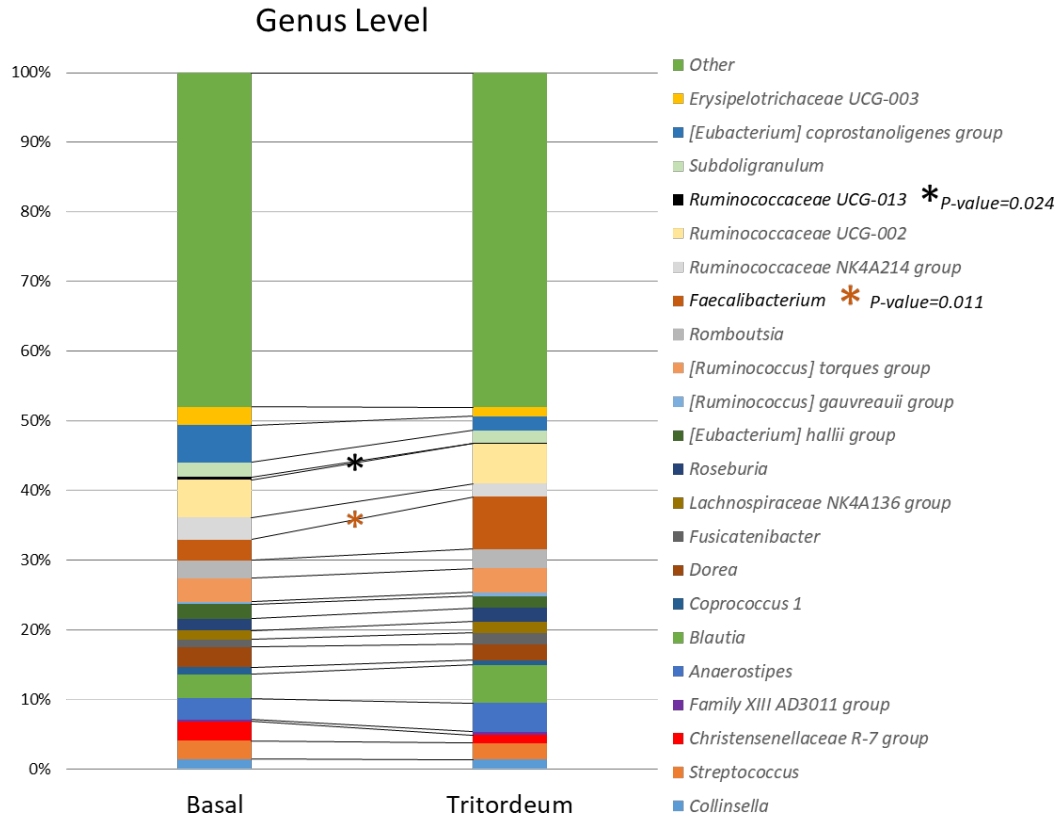


Figure 6. Summary taxa of Basal and Tritordeum phases at genus level. Graphs represented show the bar plots with the relative abundance of each taxonomic group in percentage within each phase, Basal and Tritordeum, at genus level. Other corresponds to the sum of unassigned taxa with the Silva 132 database and the bacterial taxa present in less 80% of the samples per each phase. Colored asterisks show significantly different taxa with *P*-value < 0.05. The statistically significant differences between each phase were tested by non-parametric Mann–Whitney U test.

Discussion

The study presented here follows a previous study on acceptance, digestibility and immunotoxicity of tritordeum, which showed that, in comparison with bread wheat, tritordeum breads were well accepted, had less gluten (49% reduction), lower levels of immunogenic gliadin epitopes (59% reduction in γ-gliadin epitopes, 77% reduction in α-gliadin epitopes) and that the levels of GIP excreted in stool samples of healthy volunteers were significantly reduced (Vaquero et al., 2018). The present work extends the previous study by examining the tolerance (gastro-intestinal health) and acceptance of tritordeum

breads by patients diagnosed with NCGS who habitually follow a GFD and to evaluate the effect of its consumption on the gut microbiota of these patients.

Considering the tolerance of tritordeum breads by the trial subjects, the analysis of GSRS questionnaires completed at the end of the trial showed that there was no difference in the global mean gastrointestinal health score from the Tritordeum diet and the Basal (gluten-free) diet. Also, at the individual level there were no significant differences in gastro-intestinal health scores from the two diet phases. Thus, over the seven-day intervention period at the level of tritordeum bread consumed (100 – 150 g daily, from four slices of bread) all the NCGS patients showed good tolerance, with no deterioration in gastro-intestinal wellbeing.

The measurement of GIP levels in stool samples after each of the two phases of the trial showed that, as expected, during the Basal phase subjects had lower GIP excretion, but that GIP excretion was observed during the Tritordeum phase, confirming the consumption of gluten. The levels of GIP detected in stool samples from the Tritordeum phase (mean 1.5 µg GIP/g feces) were similar to the levels detected in the previous study using healthy subjects, which had a mean of 1.2 and a range of 0.16 to 14.87 µg GIP/g feces and in which the subjects consumed the same quantity of tritordeum bread daily (Vaquero et al., 2018). However, two individuals showed high levels of GIP (2.5 and 2.9 µg GIP/g feces) excretion during the Basal (gluten-free) phase, indicating gluten exposure (volunteer or not) of the GFD. Values of GIP excretion for these two individuals in the Tritordeum phase were 1.1 and 1.9 µg GIP/g feces, respectively.

Considering the evaluation of acceptance of tritordeum bread in comparison with the gluten-free breads habitually consumed by the trial subjects, the Sensory Questionnaire showed that tritordeum breads were given significantly higher scores for all five of the acceptance attributes. Taste and texture in tritordeum bread were rated particularly highly in comparison with gluten-free breads and the subjects had a clear overall preference for tritordeum breads, awarding a mean Global Acceptance score of 7.2 for tritordeum versus 5.1 for gluten-free breads.

There is considerable controversy surrounding NCGS, because unlike the other gluten-related disorders CD and WA whose diagnosis is based on objective parameters, supported by accepted biomarkers, the diagnosis of NCGS is based on the presence of intestinal and extra-intestinal symptoms associated with the consumption gluten-containing foods and on negative results for CD and WA biomarkers (Catassi et al., 2015). At the pragmatic level, NCGS sufferers are diagnosed by the exclusion of CD and WA and by the disappearance of symptoms on the adoption of a GFD (Casella et al., 2018). Further, there is active discussion as to whether NCGS is a disease separate from irritable bowel syndrome (IBS), as several studies suggest that the symptoms overlap and that FODMAPS may play a major role in addition to / rather than gluten (Volta et al., 2016; Dale et al., 2019), although other studies present evidence that gluten-induced reactions are responsible for NCGS symptoms

(Capannolo et al., 2015; Di Sabatino et al., 2015) and ATIs also may play a role (Dale et al., 2019).

While the exact causal mechanism for NCGS has still to be defined, it appears evident that there exist a very significant number of people who report symptoms when they consume gluten-containing foods, with frequencies estimated between 6 and 13% in different populations (Foschia et al., 2016). These figures are related to a strong increase in the avoidance of wheat and other gluten-containing cereals over the last decade, with estimates up to 10 – 20% of populations in Westernised countries (Dale et al., 2019). The avoidance of wheat has resulted in an almost exponential increase in the adoption of GFDs (Foschia et al., 2016), although there is increasing recognition that at the nutritional level GFDs tend to be inferior to cereal-containing diets due to increased levels of fats, and reduced levels of complex carbohydrate, protein, vitamins and minerals (Vici et al., 2016; Larretxi et al., 2019). In addition, there is overwhelming evidence for the positive effects of the consumption of dietary fibre as supplied by wholegrain cereals, and it is more difficult to consume the recommended levels of dietary fibre when observing a GFD (Aune et al., 2016; Vici et al., 2016).

In contrast to CD, there is evidence that some NCGS sufferers may not need to totally exclude cereals from their diets, but may be able to tolerate a reduced gluten intake or gluten sources with reduced levels of GIP, either from the use of alternative cereal flours or by gluten modification via processes such as sourdough fermentation (Calasso et al., 2018; Gobbetti et al., 2018). This concept is supported by the results of the present study, in which a group of ten medically-diagnosed NCGS sufferers who habitually consume a GFD were able to consume 100 – 150g of tritordeum bread daily, giving an estimated daily gluten consumption of 5 – 7.5 g over a period of seven days, without any of them reporting a change in gastro-intestinal health parameters.

Finally, evidence is emerging that GFDs may negatively affect the gut microbiota, reducing beneficial genera of gut bacteria (De Palma et al., 2009; Lorenzo Pisarello et al., 2015; Haro et al., 2018). In this sense, our study has not observed global changes in the diversity of the intestinal microbiota between the Basal phase and the tritordeum phase. However, we have found a significant decrease in the relative abundance of *Ruminococcaceae UCG-013* genus and a significant increase of *Faecalibacterium* genus after consumption of tritordeum bread; both genera belonging to *Ruminococcaceae* family. Members of this bacterial family are considered specialists in the degradation of complex plant material, cellulose and hemicellulose, which are indigestible to the host. These compounds would be degraded by these bacterial groups to SCFA that the host can absorb and use as an energy source. For this reason, *Ruminococcaceae* is a family associated with maintaining intestinal health (Biddle et al., 2013) .

Even more interesting are the observed increase in the relative abundance of *Faecalibacterium* genera after seven days of consumption of tritordeum bread.

Faecalibacterium is one well-known as the butyrate producing bacteria within the human intestinal gut microbiota (Louis and Flint, 2009) . A new pathogenic mechanism involved in NCGS has recently been proposed in which butyrate producing bacteria, especially *Faecalibacterium* among others play a key role. When these bacteria are in optimal concentration, its production of butyrate stimulates the secretion of mucin for the protection of the mucus layer, in addition this metabolite serves as energy for the enterocytes and reinforces their tight junctions avoiding intestinal permeability (Leccioli et al., 2017). In this regard, there are several studies that demonstrate the existence of intestinal permeability in patients with NCGS (Sapone et al., 2011; Uhde et al., 2016; Hiippala et al., 2018). Finally, butyrate induces the production of antimicrobial peptides favouring microbial detoxification. Taken together all these functions, butyrate promotes intestinal homeostasis and immune tolerance (Leccioli et al., 2017).

In line with this, our results show how the consumption of tritordeum could be a good alternative to GFD in patients with NCGS, since it increases the abundance of *Faecalibacterium* which have an important role to strengthen the function of the intestinal barrier and relieve inflammation, characteristics that are compromised in this pathology (Sapone et al., 2011; Daulatzai, 2015; Uhde et al., 2016; Leccioli et al., 2017; Volta et al., 2017; Hiippala et al., 2018). In addition, according to the PLS-DA analysis, the microbial community of both phases tend to separate associated to differences of important SCFA-producing bacteria. Thus, we have observed that after the consumption of the tritordeum bread, besides *Faecalibacterium* genus, that increases its relative abundance significantly, other SCFA-producing bacteria show a similar trend including [*Ruminococcus*] *gauvreauii* group (Parkar et al., 2019), *Roseburia* (Louis and Flint, 2009; Biddle et al., 2013; Leccioli et al., 2017), *Clostridium sensu stricto 1*(Wang et al., 2019), *Bacteroides* (Rios-Covian et al., 2017), *Lachnoclostridium,* [*Eubacterium*] *xylanophilyum* group (Martin-Gallausiaux et al., 2018), and *Lachnospiraceae NK4A136* group (Hu et al., 2019) (Figure 5). This data is relevant since the SCFA are involvement in the preservation of the integrity of the intestinal barrier (Brenchley and Douek, 2012). To confirm those trends future studies are needed increasing the number of patients included in the study, as well as the period of consumption of tritordeum bread to confirm its positive effects on the intestinal microbiota by increasing SCFA producing bacteria.

Conclusion

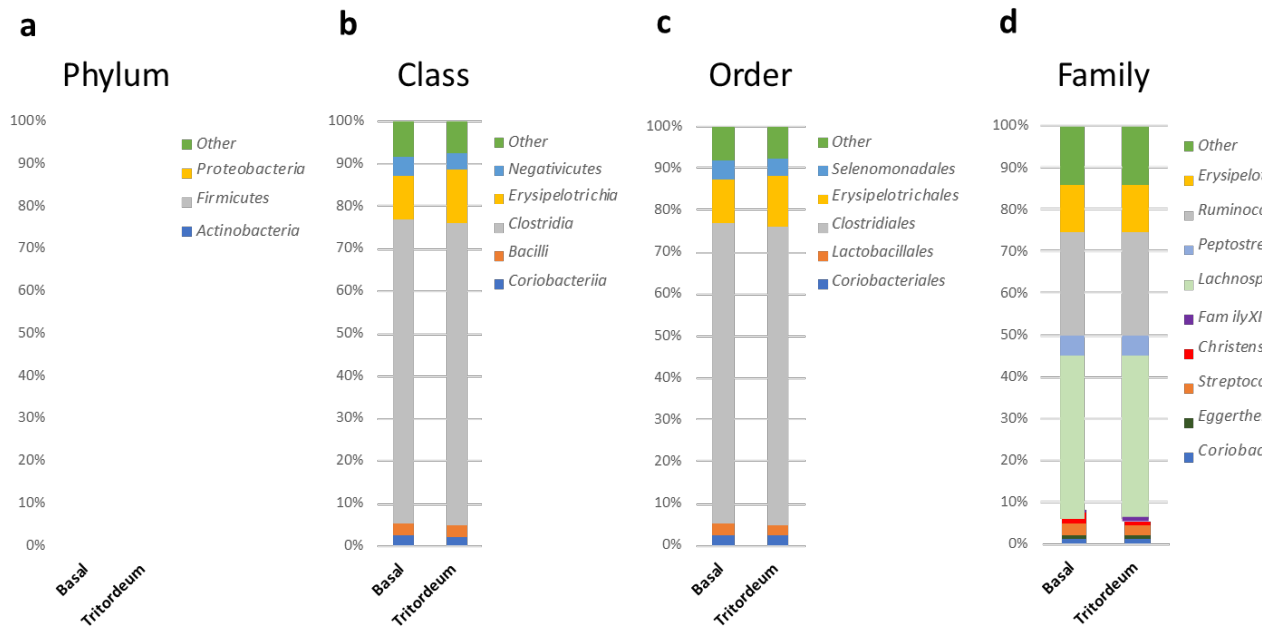
The present study shows that a group of ten diagnosed NCGS subjects who follow a GFD were able to consume tritordeum breads for a week without suffering negative effects on gastro-intestinal health. In terms of acceptability, the trial subjects showed clear preference for tritordeum breads in comparison with their usual GF bread. Tritordeum bread intake show significantly reduced levels of GIP excretion by comparison with standard wheat bread as reported in a previous study of the group, although it cannot be concluded from

the present study that this is the only factor which may have allowed its tolerance by NCGS subjects. Nevertheless, the results suggest that tritordeum may be tolerated by at least a sub-set of NCGS sufferers who do not require strict exclusion of gluten from their diet and may benefit from its better organoleptic and nutritional properties. In addition, although the global structure and composition of the intestinal microbiota of each patient was stable after tritordeum bread consumption, there was a trend to increase the relative abundance of some SCFA-producing bacteria, especially the butyrate-producing genus, *Faecalibacterium*. This result may suggest an increase of the intestinal healthy status in these patients, since the butyrate contribute to maintenance of the intestinal barrier and is known for its anti-inflammatory effects, promoting the intestinal homeostasis and immune tolerance.

Acknowledgments

The Spanish Ministry of Economy and Competitiveness (Project AGL2016-80566-P) and the European Regional Development Fund (FEDER) supported this research. The technical assistance of Ana García is acknowledged. We thank Dr. Paul Lazzeri for his review of the manuscript. We also thank the generous volunteer subjects who enrolled in the study.

Supplementary material



Supplementary Figure 1. Summary taxa of Basal and Tritordeum phases**.** Graphs represented show the bar plots with the relative abundance of each taxonomic group in percentage within each phase, Basal and Tritordeum, a) at phylum level, b) at class level, c) At order level and d) at family level. Other corresponds to the sum of unassigned taxa with the Silva 132 database and the bacterial taxa that were present in less 80% of the samples per each phase.

Supplementary Figure 2. Hierarchical cluster analysis (Euclidean distance and Ward clustering) and heatmap of bacterial taxa of the intestinal microbiota of the Basal and Tritordeum Phases of each patient.

References

Aune, D., Keum, N., Giovannucci, E., Fadnes, L.T., Boffetta, P., Greenwood, D.C., et al. (2016). Whole grain consumption and risk of cardiovascular disease, cancer, and all cause and cause specific mortality: systematic review and dose-response meta-analysis of prospective studies. *BMJ* 353, i2716. doi: 10.1136/bmj.i2716.

Aziz, I., Dwivedi, K., and Sanders, D.S. (2016). From coeliac disease to noncoeliac gluten sensitivity; should everyone be gluten free? Current Opinion in Gastroenterology 32(2), 120-127. doi: 10.1097/mog.0000000000000248.

Biddle, A., Stewart, L., Blanchard, J., and Leschine, S. (2013). Untangling the Genetic Basis of Fibrolytic Specialization by Lachnospiraceae and Ruminococcaceae in Diverse Gut Communities. *Diversity* 5(3), 627-640. doi: 10.3390/d5030627.

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology* 37(8), 852-857. doi: 10.1038/s41587-019-0209-9.

Brenchley, J.M., and Douek, D.C. (2012). Microbial Translocation Across the GI Tract. *Annual Review of Immunology* 30(1), 149-173. doi: 10.1146/annurev-immunol-020711-075001.

Calasso, M., Francavilla, R., Cristofori, F., De Angelis, M., and Gobbetti, M. (2018). New Protocol for Production of Reduced-Gluten Wheat Bread and Pasta and Clinical Effect in Patients with Irritable Bowel Syndrome: A randomised, Double-Blind, Cross-Over Study. *Nutrients* 10(12), 1873. doi: 10.3390/nu10121873.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., and Holmes, S.P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods* 13, 581. doi: 10.1038/nmeth.3869.

Canani, R.B., Costanzo, M.D., Leone, L., Pedata, M., Meli, R., and Calignano, A. (2011). Potential beneficial effects of butyrate in intestinal and extraintestinal diseases. *World journal of gastroenterology* 17(12), 1519-1528. doi: 10.3748/wjg.v17.i12.1519.

Capannolo, A., Viscido, A., Barkad, M.A., Valerii, G., Ciccone, F., Melideo, D., et al. (2015). Non-Celiac Gluten Sensitivity among Patients Perceiving Gluten-Related Symptoms. Digestion 92(1), 8-13. doi: 10.1159/000430090.

Casella, G., Villanacci, V., Di Bella, C., Bassotti, G., Bold, J., and Rostami, K. (2018). Non celiac gluten sensitivity and diagnostic challenges. *Gastroenterology and hepatology from bed to bench* 11(3), 197-202.

Catassi, C., Elli, L., Bonaz, B., Bouma, G., Carroccio, A., Castillejo, G., et al. (2015). Diagnosis of Non-Celiac Gluten Sensitivity (NCGS): The Salerno Experts' Criteria. Nutrients 7(6), 4966-4977. doi: 10.3390/nu7064966.

Cubero, J.I., Martin, A., Millán, T., Gómez-Cabrera, A., and de Haro, A. (1986). Tritordeum: A New Alloploid of Potential Importance as a Protein Source Crop1. *Crop Science* 26(6), 1186-1190. doi: 10.2135/cropsci1986.0011183X002600060023x.

Chong, J., Wishart, D.S., and Xia, J. (2019). Using MetaboAnalyst 4.0 for Comprehensive and Integrative Metabolomics Data Analysis. *Current Protocols in Bioinformatics* 68(1), e86. doi: 10.1002/cpbi.86.

Dale, H.F., Biesiekierski, J.R., and Lied, G.A. (2019). Non-coeliac gluten sensitivity and the spectrum of gluten-related disorders: an updated overview. *Nutrition Research Reviews* 32(1), 28-37. doi: 10.1017/S095442241800015X.

Daulatzai, M.A. (2015). Non-Celiac Gluten Sensitivity Triggers Gut Dysbiosis, Neuroinflammation, Gut-Brain Axis Dysfunction, and Vulnerability for Dementia. *CNS and Neurological Disorders Drug Targets* 14(1), 110-131. doi: 10.2174/1871527314666150202152436.

De Palma, G., Nadal, I., Collado, M.C., and Sanz, Y. (2009). Effects of a gluten-free diet on gut microbiota and immune function in healthy adult human subjects. *British Journal of Nutrition* 102(8), 1154-1160. doi: 10.1017/S0007114509371767.

Di Sabatino, A., Volta, U., Salvatore, C., Biancheri, P., Caio, G., De Giorgio, R., et al. (2015). Small Amounts of Gluten in Subjects With Suspected Nonceliac Gluten Sensitivity: A Randomized, Double-Blind, Placebo-Controlled, Cross-Over Trial. *Clinical Gastroenterology and Hepatology* 13(9), 1604-1612.e1603. doi: 10.1016/j.cgh.2015.01.029.

Foschia, M., Horstmann, S., Arendt, E.K., and Zannini, E. (2016). Nutritional therapy – Facing the gap between coeliac disease and gluten-free food. *International Journal of Food Microbiology* 239, 113-124. doi: 10.1016/j.ijfoodmicro.2016.06.014.

Gobbetti, M., Pontonio, E., Filannino, P., Rizzello, C.G., De Angelis, M., and Di Cagno, R. (2018). How to improve the gluten-free diet: The state of the art from a food science perspective. *Food Research International* 110, 22-32. doi: 10.1016/j.foodres.2017.04.010.

Haro, C., Villatoro, M., Vaquero, L., Pastor, J., Giménez, M.J., Ozuna, C.V., et al. (2018). The Dietary Intervention of Transgenic Low-Gliadin Wheat Bread in Patients with Non-Celiac Gluten Sensitivity (NCGS) Showed No Differences with Gluten Free Diet (GFD) but Provides Better Gut Microbiota Profile. *Nutrients* 10(12), 1964. doi: 10.3390/nu10121964.

Rideout, J.R., He, Y., Navas-Molina, J.A., Walters, W.A., Ursell, L.K., Gibbons, S.M., et al. (2014). Subsampled open-reference clustering creates consistent, comprehensive OTU definitions and scales to billions of sequences. *PeerJ* 2, e545. doi: 10.7717/peerj.545.

Rios-Covian, D., Salazar, N., Gueimonde, M., and de Los Reyes-Gavilan, C.G. (2017). Shaping the Metabolism of Intestinal Bacteroides Population through Diet to Improve Human Health. *Frontiers in microbiology* 8, 376-376. doi: 10.3389/fmicb.2017.00376.

Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4, e2584-e2584. doi: 10.7717/peerj.2584.

Roncoroni, L., Bascuñán, K.A., Vecchi, M., Doneda, L., Bardella, M.T., Lombardo, V., et al. (2019). Exposure to Different Amounts of Dietary Gluten in Patients with Non-Celiac Gluten Sensitivity (NCGS): An Exploratory Study. *Nutrients* 11(1), 136. doi: 10.3390/nu11010136.

Santos-Marcos, J.A., Rangel-Zuñiga, O.A., Jimenez-Lucena, R., Quintana-Navarro, G.M., Garcia-Carpintero, S., Malagon, M.M., et al. (2018). Influence of gender and menopausal status on gut microbiota. *Maturitas* 116, 43-53. doi: 10.1016/j.maturitas.2018.07.008.

Sapone, A., Bai, J.C., Ciacci, C., Dolinsek, J., Green, P.H.R., Hadjivassiliou, M., et al. (2012). Spectrum of gluten-related disorders: consensus on new nomenclature and classification. *BMC Medicine* 10(1), 13. doi: 10.1186/1741-7015-10-13.

Sapone, A., Lammers, K.M., Casolaro, V., Cammarota, M., Giuliano, M.T., De Rosa, M., et al. (2011). Divergence of gut permeability and mucosal immune gene expression in two gluten-associated conditions: celiac disease and gluten sensitivity. *BMC Medicine* 9(1), 23. doi: 10.1186/1741-7015-9-23.

Shewry, P.R., and Hey, S.J. (2015). The contribution of wheat to human diet and health. *Food and energy security* 4(3), 178-202. doi: 10.1002/fes3.64.

Tatham, A.S., and Shewry, P.R. (2008). Allergens to wheat and related cereals. *Clinical and Experimental Allergy* 38(11), 1712-1726. doi: 10.1111/j.1365-2222.2008.03101.x.

Uhde, M., Ajamian, M., Caio, G., De Giorgio, R., Indart, A., Green, P.H., et al. (2016). Intestinal cell damage and systemic immune activation in individuals reporting sensitivity to wheat in the absence of coeliac disease. *Gut* 65(12), 1930-1937. doi: 10.1136/gutjnl-2016-311964.

Vaquero, L., Comino, I., Vivas, S., Rodríguez-Martín, L., Giménez, M.J., Pastor, J., et al. (2018). Tritordeum: a novel cereal for food processing with good acceptability and significant reduction in gluten immunogenic peptides in comparison with wheat. *Journal of the Science of Food and Agriculture* 98(6), 2201-2209. doi: 10.1002/jsfa.8705.

Vici, G., Belli, L., Biondi, M., and Polzonetti, V. (2016). Gluten free diet and nutrient deficiencies: A review. *Clinical Nutrition* 35(6), 1236-1241. doi: 10.1016/j.clnu.2016.05.002.

Volta, U., Caio, G., Karunaratne, T.B., Alaedini, A., and De Giorgio, R. (2017). Non-coeliac gluten/wheat sensitivity: advances in knowledge and relevant questions. *Expert Review of Gastroenterology and Hepatology* 11(1), 9-18. doi: 10.1080/17474124.2017.1260003.

Volta, U., Pinto-Sanchez, M.I., Boschetti, E., Caio, G., De Giorgio, R., and Verdu, E.F. (2016). Dietary Triggers in Irritable Bowel Syndrome: Is There a Role for Gluten? *Journal of neurogastroenterology and motility* 22(4), 547-557. doi: 10.5056/jnm16069.

Wang, Y., Leong, L.E.X., Keating, R.L., Kanno, T., Abell, G.C.J., Mobegi, F.M., et al. (2019). Opportunistic bacteria confer the ability to ferment prebiotic starch in the adult cystic fibrosis gut. *Gut Microbes* 10(3), 367-381. doi: 10.1080/19490976.2018.1534512.

Yan, H., and Ajuwon, K.M. (2017). Butyrate modifies intestinal barrier function in IPEC-J2 cells through a selective upregulation of tight junction proteins and activation of the Akt signaling pathway. *PloS one* 12(6), e0179586-e0179586. doi: 10.1371/journal.pone.0179586.

CHAPTER 3

Stimulatory Response of Celiac Disease Peripheral Blood Mononuclear Cells (PBMCs) induced by RNAi wheat lines differing in the grain proteins composition

Stimulatory Response of Celiac Disease Peripheral Blood Mononuclear Cells (PBMCs) induced by RNAi wheat lines differing in the grain proteins composition

Abstract

Wheat gluten proteins are responsible of the bread-making properties of the dough but also of triggering important gastrointestinal disorders. Celiac disease (CD) affects about 1% of population in western countries. The only treatment available is a strict avoidance of gluten in the diet. RNAi is an excellent approach for the downregulation of genes coding for immunogenic proteins related to celiac disease, providing an alternative for the development of cereals suitable for CD patients. In the present work, we report a comparative study of the stimulatory capacity of seven low-gluten RNAi lines differing in the grain gluten and non-gluten proteins composition, relevant for CD and other gluten pathologies. PBMCs of 35 patients with active CD were included in this study to assess stimulatory response induced by proteins extract from the RNAi lines. Analysis of the proliferative response and INF-γ release of PBMCs demonstrated impaired stimulation in response to all RNAi lines. The lower response was provided by lines with very low content of α- and γ-gliadins, and low or almost devoid of DQ2.5 and p31-43 α-gliadin epitopes. The non-gluten protein seems not to play a key role in PBMCs stimulation.

Introduction

As basic staple foods, cereal grains substantially contribute to the dietary intake levels of energy, protein and fiber in the human diet. Wheat grains contain around 9-15% of protein, and derived wheat flour is processed into a great variety of food products that humans consume on a daily basis, such as bread, cakes, noodles, biscuits, etc. Among the seed storage proteins (SSP) of wheat grain, gluten is the most conspicuous, accounting for about a 70-80% of total protein. However, gluten is not a single protein, but a complex mixture of different proteins that accumulate during the grain development, which ultimately provides to wheat flour with unique physicochemical attributes which are a key factor in the baking quality.

Wheat gluten features outstanding technological properties and nutritional benefits, but the intake of these proteins has also been related to the triggering of certain pathologies. Gluten-related allergies and intolerances include CD, non-celiac wheat/gluten sensitivity (NCWS) (Sapone et al., 2011; Fasano et al., 2015), wheat allergy, baker's asthma, and wheat-dependent exercise-induced anaphylaxis (WDEIA) (Palosuo et al., 2001). CD is the best-known and most studied pathology, affecting adults and children with an estimated

prevalence about 1% in worldwide population (Choung et al., 2017; Lebwohl et al., 2018). In contrast, NCWS is not well understood, with symptoms overlapping with those of CD and irritable bowel syndrome (IBS) (De Giorgio et al., 2016). Prevalence of NCWS is difficult of assess as, in contrast to CD, there is not reliable immunogenic, genetics, or molecular biomarkers, and diagnosis is made by exclusion of other gut-related pathologies (Dale et al., 2019). It is not clear whether the causal agent of NCWS is gluten. In a recent study, a large proportion of patients self-reporting NCWS were instead sensitive to FODMAPs (fermentable oligosaccharides, monosaccharides and disaccharides and polyols) (Fasano et al., 2015; Skodje et al., 2018). Interest has been focused in the non-gluten components of wheat, a group of proteins that include the amylase trypsin inhibitors (ATIs), which seems to activate the innate immune system (Junker et al., 2012), and therefore they were proposed to have a role in the NCGS.

In the case of CD, gluten is a strong environmental factor, but it also has a genetic and immunological component related to human leukocyte antigen (HLA)-DQ2 and HLA-DQ8 (Tjon et al., 2010). Due to their high content of proline and glutamine, gluten proteins – also known as "prolamins" -, are resistant to their complete digestion in the human digestive tract. Thus, peptides produced as the result of the partial digestion of prolamins induce an autoimmune mediated disorder that leads into small intestine inflammation, malabsorption, and villous atrophy in patients suffering from CD. A two-signal model is currently the most accepted to explain CD (Brandtzaeg, 2006). According to this model, certain gluten peptides trigger an innate immune response followed by a secondary antigen-specific adaptive response. The result of the innate response is the increasing of the permeability of epithelial barrier, allowing gluten peptides to reach the lamina propria (Maiuri et al., 2000). The most studied innate response activator peptide is known as p31-43, occurring in α-gliadins (Maiuri et al., 2003). The autoimmune response is strongly enhanced as consequence of deamination of glutamine residues present in gluten peptides by tissue transglutaminase 2 (tTG2) in the intestinal mucosa. Deamidated peptides are capable of binding to HLA-DQ2 and HLA-DQ8 molecules present in antigen presenting cells (APC)(Tollefsen et al., 2006b), stimulating T-cells that release pro-inflammatory cytokines such as interferon-γ, TNF-α and IL-2. These cytokines damage the enterocytes and produce the intestinal lesions typical for CD (Holtmeier and Caspary, 2006). Although several CD epitopes are found in the glutenin fraction of gluten, the majority of the immunogenic CD epitopes are found in the gliadin fraction of gluten (Arentz–Hansen et al., 2002). Among the gliadins, the α-gliadins have the strongest immunogenicity, and the α-gliadin 33-mer is the main immunodominant toxic peptide in celiac patients. This peptide is present in the N-terminal repetitive region of α -gliadins and contains six overlapping copies of three different DQ2-restricted T-cell epitopes with highly stimulatory properties (Shan et al., 2002).

The only treatment available for CD patients is a lifelong strict gluten-free diet. However, complete avoiding of gluten from the diet is arduous, as these proteins are widely used in

the food industry and added to many foodstuffs which may apparently do not contain them. Besides, gluten-free products tend to be less healthy than those containing gluten, as high amounts of fat and sugar are involved in their production with the aim to provide gluten-free (GF) products with a texture that mimic the viscoelastic properties of gluten proteins (Vici et al., 2016). For this reason, the development of wheat varieties with reduced immunogenic profiles should be considered an excellent ingredient to improve the diet of patients with CD, as well as for those with NCGS. In addition, these wheat varieties could undoubtedly also apply to the general population, in particular for those who, for whatever reasons, want to reduce the intake of gluten. In fact, in a recent study, a beneficial effect of low-gluten diet was reported in comparison to high-gluten (Hansen et al., 2018). Results showed that low-gluten diet changed the gut microbiome of participants, reduced their gastrointestinal discomfort, and resulted in a small weight loss.

One promising approach to reduce gluten content and immunogenicity is the downregulation of immunodominant peptides by interference RNA (RNAi) (Hammond et al., 2000; Bernstein et al., 2001). In previous works, we implemented this technology to reduce the expression of specific gliadin fractions (Gil-Humanes et al., 2008) as well as all three gliadin fractions (Gil-Humanes et al., 2010; Barro et al., 2016) in bread wheat. Protein extracts from RNAi lines with the three gliadin fractions regulated downwards showed a pronounced reduction in T-Cell response when tested *in vitro* (Gil-Humanes et al., 2010). However, the down-regulation of both specific and all gliadin fractions provided a compensatory effect with other protein fractions (Pistón et al., 2013; García-Molina et al., 2017). This is particularly important when non-gluten proteins (NGP) are used for this compensation as metabolic proteins and chloroform/methanol soluble proteins (CM-like), such as the α-amylase/trypsin inhibitor family, β-amylase and serpins, were related to wheat allergens (Volta et al., 2013).

The presents work was designed to study the *in vitro* response of peripheral blood mononuclear cells (PBMCs) from CD patients to different RNAi lines differing in the gluten, and NGP composition. Modifications in wheat grain protein composition by RNAi led to a reduction of CD-related epitopes of the most immunogenic fractions. The findings of this research may be useful to establish the optimal protein composition to pursue the maximum reduction of the immunogenic potential for CD for the development of new wheat varieties.

Material and methods

Plant Material

Seven transgenic lines derived from bread wheat cv. Bobwhite (BW208) and their correspondent wild type were used in this study (Table 1). Transgenic lines and transformation vectors used were previously reported (Gil-Humanes et al., 2010; Pistón et al., 2011; Pistón et al., 2013; Barro et al., 2016). Lines were obtained using combinations of different RNAi fragments designed to target different gliadin fractions; ω-, α- and γ-gliadin, and LMW fraction of glutenins. All silencing fragment were expressed under a D-hordein endosperm-specific promoter (Pistón et al., 2008).

Line	Plasmid 1	Plasmid 2	Prolamin target
BW208	NA	NA	NA
D623	pghpg8.1	NA	γ-gliadin
D783	pDhp_ω/α	NA	ω-, α-, and γ-gliadin
E82	pghpg8.1	pDhp_ω/α	ω-, α-, and γ-gliadin
H320	pDhp_α/βZR	pDhp_ω4ZR	ω- and α-gliadin
H754	pDhp_ω8ZR	NA	ω-gliadin
H811	pDhp_α/βZR	pDhp_ω8ZR	LMW, ω- and α-gliadin
I17	pDhp_α/βZR	NA	α-gliadin

Table 1. RNAi lines obtained from the wild type BW208 and their corresponding transformation vectors and prolamin silencing targets. NA, non-applicable.

Reversed-phase high-performance liquid chromatography (RP-HPLC) quantification of SSP proteins

Gliadins and glutenins from 100 mg of flour from three biological replicates were extracted stepwise. Both fractions were quantified by RP-HPLC following the protocol as reported (Pistón et al., 2011).

Total protein and Non-gluten proteins quantification

The protein content of whole flour was calculated from the Kjeldahl nitrogen content (%N x 5.7) according to the standard ICC method no. 105/2 (ICC, 1994). The non-gluten proteins (NGP), expressed in percentage of dried weight (% DW), were calculated as follow: [Total protein in % – (Prolamin content in μg/mg*10)/(100 – moisture in %)].

Liquid Chromatography-Tandem Mass Spectrometry analysis

Total protein extraction and pepsin-trypsin digestion of samples was carried out from 2g of flour following a protocol previously reported (Barro et al., 2016), with two exceptions: a greater ratio flour/extraction volume, and centrifuge filtration of extracts (Corning 45 µm Nylon tubes) prior to enzymatic digestion. Protein digests (1.5 µg) were analyzed after being cleaned with a SEP-PAK C18 cartridge (Waters, Milford, MA) by one-dimensional nanoscale liquid chromatography- tandem mass spectrometry (LC-MS/MS) on an Eksigent NanoLC-1D plus (AB SCIEX) coupled to a 5600 Triple TOF® mass spectrometer (AB SCIEX) equipped with Acclaim PepMap 100, 100 lm 9 2 cm (Thermo Fisher Scientific, Waltham, MA) precolumn and NanoACQUITY UPLC1.7 lm BEH130 C18, 75 lm 9 150 mm (Waters) HPLC column. For proteomic analysis resulting data were searched against NCBI protein database for *Triticum aestivum* species without any enzyme restriction. Searches were conducted using Mascot Server 2.4 (Matrix Science, London, UK), with a peptide mass tolerance of 25 ppm and fragment mass tolerance of 0.05 Da. Only peptides with scores higher than 20 were extracted for further analyses. CD Epitopes content were determined using Blastp to search against celiac disease epitopes described by Sollid and collaborators (Sollid et al., 2012) and the α-gliadin peptide 31–43 in peptides longer than eight amino acids identified by LC-MS/MS analysis.

Gluten content determination by competitive enzyme-linked immunosorbent assay (ELISA)

Gluten content of whole flour was measured by G12 monoclonal antibody (moAb) as described previously (Barro et al., 2016). Each sample were measured in triplicated. Results were expressed in parts per million (ppm) in dry matter.

Peripheral Blood Mononuclear Cells Proliferation and IFN-γ Production analysis

Patients with active CD on a gluten-containing diet (n=35) were included in this study. The diagnosis of CD was primarily determined by serological screening tests and finally confirmed by biopsy of the small intestine. The mucosal specimens were graded independently according to the Marsh–Oberhuber classification (Marsh, 1992; Oberhuber et al., 1999). Subjects were prospectively screened for CD using antiendomysial antibodies (AAEMs), anti-tissue transglutaminase antibodies (AATGs), and CD-specific HLA (human leukocyte antigen) typing (Table S2). The local Ethics Committee of the Hospital 'Virgen de las Nieves' (Granada, Spain) approved the study protocol. Written consent was obtained from parents or legal guardians of children.

PBMCs were isolated from 6 mL of heparinized blood by Histopaque gradient centrifugation (Sigma Aldrich, Madrid, Spain) and cultured at a density of 1 × 106 cells/ml in RPMI-1640 culture medium (Gibco,Thermo Scientific, Madrid, Spain) supplemented with 10% fetal

bovine serum (GIBCO-Invitrogen Ltd), 1% penicillin-streptomycin, and 0.1% gentamicin (Sigma-Aldrich).

The above described PT-digested protein extracts were also used to study immunogenic potential by PBMCs proliferation assay and IFN-y release. The above described PT-digested protein extracts were also used to study immunogenic potential by PBMCs proliferation assay and IFN-y release. Rice flour and synthetic extract of 33mer peptide were used as negative and positive controls respectively. After 48 h of culture PBMCs were incubated with protein extracts from the different lines and controls (rice and blank without protein extracts added). Each experiment was carried out in duplicate. Cultures were collected separating PBMCs for cells proliferation studies and supernatants for IFN-y analysis after 24 h of stimulation. Supernatants from the PBMCs culture were stored at -80 °C until IFN-γ determination were carried out using a commercial ELISA kit (Thermo Scientific, Madrid, Spain) in accordance with the manufacturer's instructions. Standards were run on each plate. Assay sensitivity was less than 2 pg/mL.

Cells proliferation was determinate by the ELISA 5-bromo-2-deoxyuridine (BrdU) cell proliferation test (Millipore Chemicon, California, USA). Proliferative responses of PBMCs were defined as a stimulatory index (SI), this variable represent the specific proliferation of a sample as the mean absorbance at 450 nm after stimulation divided by the background proliferation (mean absorbance of PBMCs exposed to the culture medium alone).

Statistics

Statistical software R version 3.5.1 (Ihaka and Gentleman, 1996) was used for data analysis and some plots. Analysis of variance (ANOVA) followed by two-tailed Dunnett's test for mean multiple comparisons was used for differences between lines. Normal distribution and homogeneity of variance were previously tested by Shapiro–Wilk normality test and Levene test respectively. Pearson's R was used to determine data correlation. Figures were drawn using the Microsoft Excel and PowerPoint software (Microsoft Corporation). The libraries FactoMineR and factoextra were used for PCA analysis and graphical output, respectively.

For cells proliferation and IFN-γ assays each experiment was carried out in duplicate on separate days. The data is expressed as mean and SD. All statistical analyses were performed with the STATGRAPHICS Centurion XVI program. The analysis of variance (ANOVA), followed by the Tukey test for mean multiple comparison, was used.

In this study, *P* values lower than 0.05 were considered significant

Results

Grain protein composition in RNAi lines

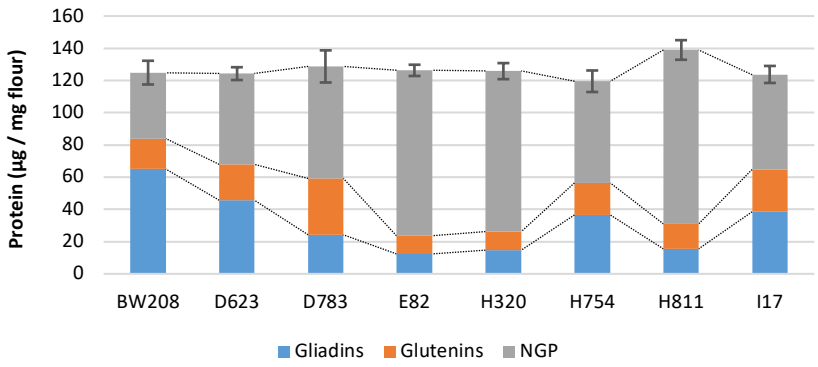
The contents of the different gluten and NGP fractions of the wild type BW208 and the RNAi lines are shown in Figure 1. All RNAi lines present reduction of the gliadin fraction, while the range of variation for glutenins depends on the line analyzed. The strongest reduction for gliadins corresponds to lines E82, H320, and H811 (Figure 1a), which all three are targeted with two plasmids (Table 1). Total glutenins of lines D623, H754 and H811 do not present significant change with respect to the BW208 wild type line. In contrast two lines, D783 and I17, show upregulation of the glutenin fraction whereas, in two other lines, E82 and H320, there is a reduction of total glutenins. As shown in Figure 1a, the NGP fraction is upregulated for all RNAi lines, and particularly for lines E82, H320, and H811. The reduction in the gliadin fraction of the grains is compensated through the overexpression of other proteins so that the total protein remains unaffected for all lines except for line H811, in which the total protein content is significantly affected in comparison to that of the BW208 wild type.

Focusing on the gluten fractions, which include gliadins and glutenins (Figure 1b), all RNAi lines show significant changes in some of the protein fractions in comparison to the BW208 wild type, irrespective of the fragment used for silencing. Lines D623, H754 and I17 hold RNAi vectors aimed to silence a single gliadin fraction (γ-, ω- and α-gliadins, respectively), but only line D623 shows a reduction only on its specific target (Figure 1b). Besides of downregulation of their specific targets line I17, targeting α-gliadins, also presents lower amount of γ-gliadins, and line H754, targeting ω-gliadins, also showed a decrease of the α- and γ-gliadin fractions. E82 is the line with the highest reduction of gliadins, followed by lines H320 and H811, both with α-and ω-gliadins as targets, but with strong reduction in the γ-gliadin fraction. Line D783 share one construct with line E82, but the latest has an additional construct to target more precisely the γ-gliadin fraction. In comparison to D783, E82 line has a strong reduction in the γ- and α-gliadins, suggesting a synergic effect when more than one gliadin fraction is used as silencing target.

Within the glutenins, the HMW fraction is significantly up-regulated in comparison to that of the BW208 wild type in lines D783, H754, H811 and I17, and not significantly affected in the other three RNAi lines. HMW glutenins are not included as target in any of the RNAi vectors and do not present down-regulation in any of the RNAi lines. However, probably due to the grain protein compensation system, there is an overexpression of HMW in most lines, particularly in line D783, with up to a 3-fold increment. In contrast, the LMW fraction is significantly affected in all but one RNAi lines; in lines D623 and D783 the LMW are up-regulated while in lines E82, H320, H754, and H811 are strongly down-regulated (Figure 1b).

Overall, E82, H320 and H811 are the lines with a higher reduction in prolamins, being the differences among them in the stronger reduction in γ-gliadins for E82, and the stronger up-regulation of HMW in line H811.

a



b

		BW208	D623	D783	E82	H320	H754	H811	I17
Gliadins	Omega	10.8	13.2 *	6.8 ***	6.8 ***	7.0 ***	8.3 *	7.7 **	10.8 ns
	Alpha	27.5	29.6 ns	12.9 ***	5.5 ***	5.0 ***	17.7 ***	4.6 ***	9.5 ***
	Gamma	26.7	2.8 ***	4.6 ***	0.1 ***	2.6 ***	10.7 ***	3.2 ***	18.2 ***
Glutenins	HMW	6.3	6.8 ns	19.4 ***	8.8 ns	8.8 ns	15.5 ***	13.4 ***	13.3 ***
	LMW	12.6	15.6 *	15.6 *	2.5 ***	3.0 ***	4.3 ***	2.2 ***	13.1 ns

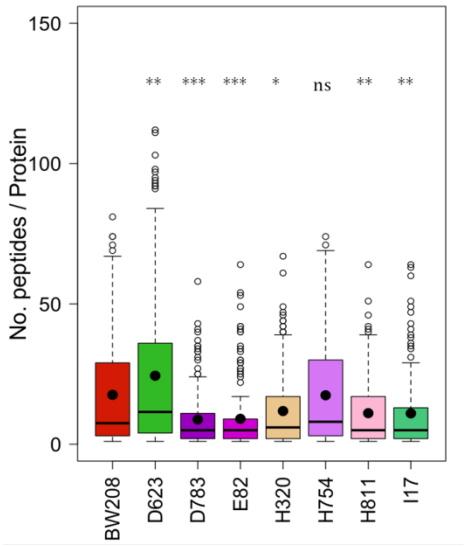
Figure 1. Grain protein composition (µg protein /mg flour) of RNAi transgenic lines and wild type (BW208). A) Protein distribution of three main protein groups in the wheat grain, B) Protein distribution of prolamin fractions. NGP, Non-Gluten Proteins; HMW, High Molecular Weight glutenin subunits; LMW, Low Molecular Weight glutenin subunits. Dunnett multiple comparison of means with BW208 wild type line; ns = no significant; * $P < 0.05$, ** $P < 0.01$; *** $P < 0.001$.

Analysis of pepsin and trypsin protein extracts digestion (PT-digested)

Total protein was extracted from flour of the six RNAi wheat lines and the BW208 wild type line and then subjected to pepsin and trypsin digestion (PT-digested). Resulting fragments were analyzed by LC-MS/MS and identified through a *Triticum aestivum* species restricted search in the National Center for Biotechnology Information (NCBI) database. The silencing of specific prolamin fractions by RNAi resulted in significant differences in the number of peptides per protein identified in PT-digested flour for all RNAi lines except H754 (Figure 2a). The line D623 showed a higher number of peptides per protein than the BW208 wild type, whereas all other lines but H754 provided significant lower number of peptides per protein. A variation in the number of peptides per protein reflects differences in the protein composition after silencing by RNAi. In this regard, lines H320, H811 and I17 showed the

lowest number of α-gliadin peptides in comparison to the BW208 wild type (Figure 2b). In contrast, the line D623 had an important increase in α-gliadin peptides. In general, lines E82, D783, and H811 showed reduction for the peptides of all three gliadin fraction. Compensatory effects described above for HPLC data is also reflected in the number of peptides per protein in the glutenins and in the NGP fractions. For example, the number of identified peptides corresponding to ATIs, globulins, serpins, triticins, and other NGP are notably increased in some RNAi lines (Figure 2b). Particularly interesting is the increment in serpin peptides for all RNAi lines.

a



b

	BW208	D623	D783	E82	H320	H754	H811	I17
Gluten proteins								
Alpha	1006	1860	126	28	0	805	7	9
Gamma	469	304	195	29	595	631	279	262
Omega	259	168	23	10	168	251	72	279
HMW	11	13	34	166	38	13	35	5
LMW	137	327	140	79	79	66	47	72
NGPs								
ATIs	24	11	40	36	39	44	32	27
Globulins	81	75	87	96	90	85	122	93
Serpins	47	76	164	221	188	138	180	92
Tritins	10	8	23	24	18	15	22	14
Triticins	8	10	25	60	27	18	28	17
Other NGP	178	130	232	280	270	220	219	173

Figure 2. Average number of peptides per protein and genotype identified by LC/MS-MS (A). Number of protein peptides identified corresponding to the different gluten and non-gluten protein fractions (B). NGPs, Non-Gluten Proteins; ATIs, amylase trypsin inhibitors; HMW, High Molecular Weight glutenin subunits; LMW, Low Molecular Weight glutenin subunits. Dunnett multiple comparison of means with BW208 wild type line; ns = no significant; * $P < 0.05$, ** $P < 0.01$; *** $P < 0.001$.

Peptides were searched for the presence of DQ restricted epitopes, related with adaptive immune response, and for the p31-43 peptide present in α-gliadins and linked with innate immune activation (Supplementary table 1). Results of the analysis showed a notable decrease in the number of CD epitopes in four of the RNAi lines (Figure 3a). There is only one line (H754) that presents a higher number of total epitopes in comparison with the wild type. Of the epitopes identified in the wild type line, the vast majority correspond to γ- and α-gliadins epitopes. The p31-43 fragment hits were reduced in all RNAi lines except D623, and it was not found in I17, H320 and H811. The γ-gliadin related epitopes are the most abundant in all samples analyzed including the wild type line. The reduction in the number of epitopes found in this gliadin fraction ranges from 4% in line H320 up to 95% in E82 when compared with the wild type. Epitopes corresponding with α-gliadins were not identified in peptides from three of the RNAi lines (I17, E82 and H320) and were reduced more than 93% in lines D783 y H811. In contrast, they appeared notably increased in lines

D623 and H754 where they are overexpressed. Epitopes present in ω-gliadins are also reduced in all lines, with an average reduction of more than 80% in lines containing inverted repeated (IR) fragment for this target and being absent in line E82. Regarding glutenin epitopes, half of lines (D623, D783, E82 and H320) had higher content of these epitopes meanwhile the other three lines (H754, H811 y I17) present a decrease in the number of epitopes for this protein fraction (Figure 3a). Results from proteomic analysis showed that lines with a greater reduction in the total number of immunogenic CD epitopes are E82, D783, and I17, with an average reduction of 94, 88, and 62%, respectively, compared with the wild type. It is noteworthy that line E82 appeared to be completely devoid of CD epitopes from the highly immunogenic α- and ω-gliadins and containing only a 5% of γ-gliadin ones.

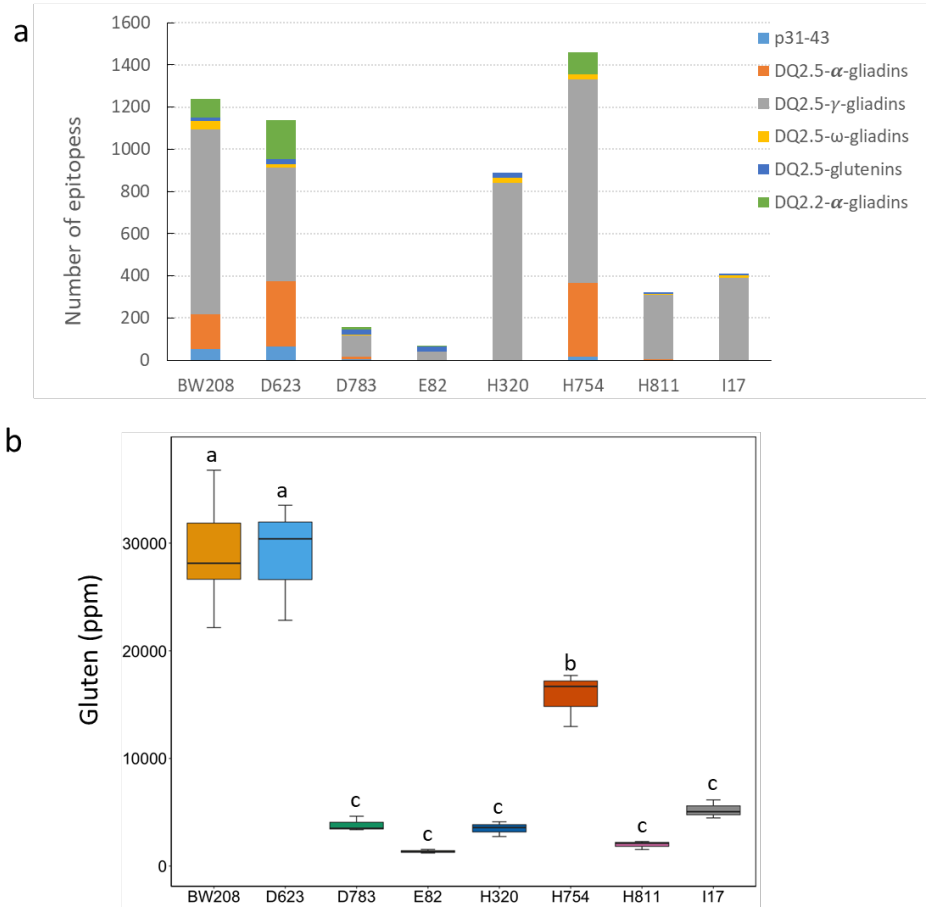


Figure 3. Number of DQ restricted epitopes identified in LC/MS-MS peptides (A). Gluten content as measured by monoclonal antibody G12 ELISA expressed in parts per million (ppm) (B). Lines with the same letter are not significant different according to the Tukey test comparison of means ($P < 0.05$).

Gluten immunogenicity by G12 moAb

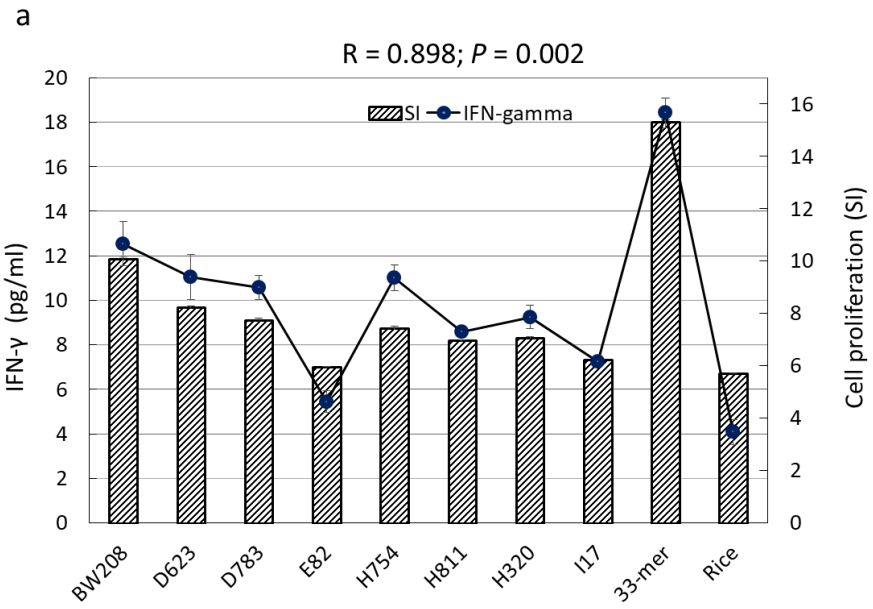
Gluten content in parts per million (ppm) of flour from all lines was determined by a competitive ELISA system with G12 moAb (Morón et al., 2008b). As showed, gluten content

(ppm) was strongly reduced in 5 of the 7 lines compared with wild type (Figure 3b), with an average reduction of 85% and a maximum reduction of 95% corresponding to line E82. RNAi line D623 showed comparative gluten content to that of the wild type while line H754 had a significant lower value in comparison to the wild type. Overall, lines E82 followed by H811 are the ones with lower immunogenic potential. The level of reduction in G12 values were strongly correlated with the content of α-gliadins (R = 0.972; $P < 0.001$), total gliadins (R = 0.888; $P = 0.003$), and total prolamin (R = 0.772; $P = 0.025$) fractions as determined by RP-HPLC. In addition, G12 also provided significant correlations with the number of DQ2.5 from α-gliadins (R = 0.845; $P = 0.007$) and p31-43 (R = 0.973; $P < 0.001$) peptides, present in α-gliadins, and with the total number of gliadin peptides (R = 0.729; $P = 0.041$).

Stimulatory response of PBMC from CD patients

Cell proliferation assay and IFN-γ response were carried out using the PT-digested flour from the RNAi lines in order to determine their ability to activate stimulatory response of PBMCs from CD patients. Immunogenic potential was tested as a measure of capacity to trigger cellular proliferation (Stimulation Index, SI) in cultures from 35 different child patients suffering from celiac disease and under gluten containing diet (Supplementary table 2), using the BW208 wild type flour and the 33-mer peptide as positive controls, and PT-digested rice flour as negative control. Also, a control without proteins extracts added in the culture cells (blank group) was included as reference values to compare the effect of peptides in cells of patients with CD, under the same conditions of cell culture. There is an excellent correlation between cell proliferation and IFN-γ release (R = 0.898; $P = 0.002$). As expected, the 33-mer peptide and rice negative control showed the maximum and minimum, respectively, of both cell proliferation and IFN-γ release. BW208 was the wheat line with a higher cell proliferation and IFN-γ values (Figure 4). As showed, changes observed in the grain prolamin fractions, as consequence of RNAi silencing, led to changes in the cell proliferation and IFN-γ release of RNAi lines. All RNAi lines have cell proliferation values (SI) lower than the BW208 wild type line, showing an average of cell proliferation reduction of about 30%. A reduction in proliferative response of PBMCs was seen in lines E82, I17 and H811, with SI values similar to that of the rice negative control (Figure 4, Table S3). In contrast, lines BW208, D623, D783, H320 and H754 showed significant higher values than the rice negative control. IFN-γ release confirmed impaired stimulatory capacity of PT-digested flour from the seven RNAi lines (Figure 5), with RNAi lines showing an average reduction of 28% with respect to BW208 wild type. Lines E82 and I17 provided the lowest values and, in the case of E82 it were not significantly different to that of the rice negative control (Figure 4, Supplementary table 4), demonstrating a decrease in the immunotoxicity of the flour by reduction of certain gluten fractions by RNAi. Cell proliferation was positively correlated with both the α-gliadin (R = 0.745; $P = 0.034$) and total gliadin (R = 0.727; $P = 0.040$) contents as determined by HPLC. In addition, cell proliferation showed significant correlation with gluten content provided by G12 MoAb (R = 0.768; $P = 0.026$)

and with the number of p31-43 (R = 0.754; *P* = 0.030) and DQ2.5 ω-gliadins (R = 0.741; *P* = 0.035) epitopes detected in mass spectrometry peptides.



b

Cell proliferation (SI)

	Homogeneous Groups
Control	a
E82	ab
IL7	abc
H811	abcd
H320	bcd
H754	cd
D783	d
D623	d
Bw208	e
33-mer	f

c

IFN-γ production

	Homogeneous Groups
Control	a
E82	a
IL7	b
H811	c
H320	c
D783	cd
D623	d
H754	de
Bw208	e
33-mer	f

Figure 4. (a) Immunogenicity of different RNAi wheat lines. Bars represent the proliferative responses of PBMCs to PT-digested protein extracts from the wild type (BW208) and RNAi wheat lines defined as a stimulatory index (SI). Line and dots represent the Interferon-γ (IFN-γ) production by PBMCs with PT-digested from the wild type (BW208) and RNAi wheat lines. Results represent the mean of 35 patients ± Standard Deviation (SD). (b) Cell proliferation and (c) IFN-γ production homogeneous groups according to Tukey HSD multiple range test. Different letters denote significant differences (P < 0.05).

Our data confirm notable differences between the RNAi lines for cell proliferation and IFN-γ release. Overall, the RNAi lines E82 and 17 were highly inefficient to stimulate the immunogenic response by cell proliferation and IFN-γ release. Both lines do not differ significantly from rice negative control (Figure 4, Supplementary table 3 and 4) indicating that the reduction of the gliadin fractions in these lines, together with the reduction of LMW and the increment of HMW and NGP, result in a significant decrease in the immunogenic potential of their protein extracts.

The relationships described above among all the parameters tested (quantification by HPLC of protein fractions, results of proteomic analysis and proliferation of PMBC) are well reflected in the PCA analysis (Figure 5). The first two dimensions represent 80% of the variation of the set of variables in the seven lines studied, and show that all variables related to immunogenicity have a positive value in the first axis; that is, they vary in the same direction as the proliferation values of PBMCs and IFN-γ. These last two have an almost null value for the third dimension, which represents 9% of the variation, which

implies that its variation is well explained only with the first two axes. All variables related to NGPs and HMW glutenins vary in the opposite direction on the first axis, suggesting that they do not contribute to the immunotoxicity of the lines for celiac patients.

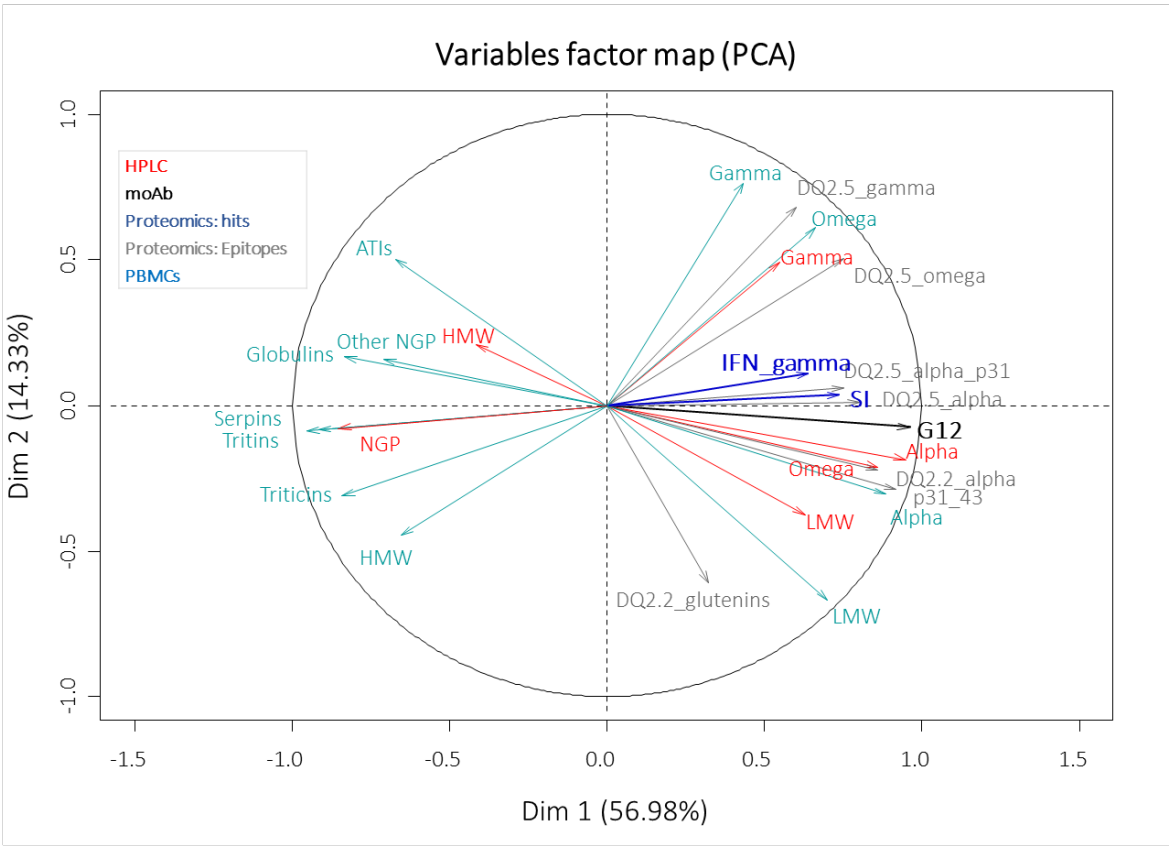


Figure 5. PCA analysis of the variables measured by HPLC, Proteomics, PBMCs and moAb G12

Discussion

RNAi technology has demonstrated being a powerful approach for the down-regulation of different gluten genes, whose proteins are responsible of CD and other gluten-related pathologies. RNAi has been used to down-regulate the expression of specific gliadin genes; γ-, ω-, and α-gliadins (Gil-Humanes et al., 2008; Barro et al., 2016; Altenbach et al., 2019), or even gliadin genes from more than one family at once (Gil-Humanes et al., 2010; Wen et al., 2012; Barro et al., 2016). The proteome of these lines was studied in detail showing a compensatory mechanism to fill the gap left by the gliadins being silenced (García-Molina et al., 2017; Altenbach et al., 2019). This compensatory mechanism is operating on other gluten or NGP, and therefore the glutenins (HMW and LMW), and NGP like serpins, ATIs, globulins, or triticins are increased besides the silencing of gliadins. Therefore, the RNAi down-regulation has a direct effect on the composition and proportion of the gluten and non-gluten fractions. In this work, we report a comparative study of the stimulatory

response of RNAi lines, targeted at different gliadin or glutenin fractions, and differing in the grain protein composition relevant for CD and other gluten pathologies.

This set of RNAi lines (Table 1) were produced by the expression of one or two RNAi silencing fragment under the control of a D-Hordein endosperm specific promoter. Consequently, specific gliadin fractions as well as all three gliadin families were strongly down-regulated. Although downregulation of specific gliadins is observed, stronger downregulation of the targeted gliadin groups is provided when RNAi constructs targeting different groups work together in the same line. This pleiotropic effect is clear in lines E82 and D783 which share one construct targeting all gliadins groups. However, line E82 also includes a construct (pghpg8.1) designed to target specifically the γ-gliadins, and E82 presents a higher reduction of the γ-gliadins than D783. The specificity of pghpg8.1 is reflected in line D623 where silencing only occurs in γ-gliadins (Figure 1b). This fact suggests a synergic effect when more than one gliadin sequence is used as silencing target. The same occur in lines H811 and I17, both contain a construct directed to α-gliadins, but this group of proteins appear more reduced in line H811 (Figure 1b), which has an additional construct targeting also ω-gliadins (Table 1).

Silencing of specific gluten proteins by RNAi provided a compensatory mechanism with other proteins, resulting that total protein content does not present significant changes in the RNAi lines in comparison to the wild type, with the exception of line H811. This compensatory mechanism involves mainly the NGP and HMW glutenins. The latest are major determinants of bread-making quality and the higher proportion of this fraction, present in some lines, could indicate reasonable baking quality. Gil-Humanes et al. (Gil-Humanes et al., 2014a) reported lines with increased HMW as consequence of RNAi silencing, providing flours with increased stability and better tolerance to over-mixing. In agreement with this, Altenbach et al. (Altenbach et al., 2019) reported an increment of the HMW subunits when ω-1,2 gliadin genes were down-regulated by RNAi, leading to an increment in the mixing time and tolerance. Up regulation of NGP is occurring in all lines independently of the gluten fraction silenced. The increment of NGP, particularly triticins (Gil-Humanes et al., 2011), let to wheat lines with improved nutritional properties since its lysine content is significantly higher than that of normal flour (Gil-Humanes et al., 2014a). The NGP includes metabolic and structural proteins, CM-like and serpin proteins, which has been related not only to wheat allergy (Larré et al., 2011; Matsuo et al., 2015) but also as novel target antigens in CD humoral response (Huebener et al., 2015). The compensatory mechanism was corroborated with the analysis of PT-digested proteins from the flour of RNAi lines. An increment in the number of glutenin and NGP peptides was found in the RNAi lines. This increment is particularly important for ATIs, serpins and triticins, which corroborate previous observations using 2-D gel electrophoresis and LC-MS/MS analysis (García-Molina et al., 2017; Altenbach et al., 2019).

Gliadin fraction of gluten seems to play a major role in CD as the most immunogenic epitopes are on this proline and glutamine rich proteins (Arentz–Hansen et al., 2002). PT-

digested peptides from RNAi lines were searched for the presence of relevant T-cell epitopes recognized by CD4+ T cells and binding to human leukocyte antigens (HLA)-DQ. DQ2.5 restricted epitopes have higher risk for CD (Sollid et al., 2012) because of the ability of HLA-DQ2.5 molecules to form stable complexes with a large gluten peptides repertoire (Fallang et al., 2009). DQ2.5 epitopes encoded by α-gliadins were not found in three of the RNAi lines, while those DQ2.5 epitopes encoded by γ-gliadins were strongly reduced in four lines. However, DQ2.2 and p31-43 α-gliadin epitopes were found in line E82 at very low frequency. Lines devoiding of α-gliadin epitopes still contain α-gliadin proteins as determined by RP-HPLC (Figure 1b) and corroborated by mass spectrometry analysis (Figure 2b). As reported by Ozuna et al. (Ozuna et al., 2015), there is a significant number of α-gliadin sequences devoid of immunogenic epitopes, and this could explain in part the observed discrepancy between RP-HPLC and the number of epitopes detected by mass spectrometry. On the other hand, gliadin family is a complex mix of proteins and the exact number of genes encoding those proteins is variable between genotypes and not well known yet. Therefore, this is an additional limitation of protein databases as not all protein sequences and peptides are deposited in them. Despite of the mentioned limitations, there was an excellent correlation between the α-gliadins determined by RP-HPLC and the number of α -gliadin peptides determined by mass spectrometry ($R = 0.841$; $P = 0.009$). Moreover, results by G12 were also correlated with both the content of α-gliadins as determined by RP-HPLC ($R = 0.972$; $P < 0.001$) and the number of α-gliadin epitopes as determined by mass spectrometry ($R = 0.845$; $P = 0.008$).

All lines with the exception of D623 have a significant reduction in the gluten content detected by the G12 moAb. This antibody was developed against the 33-mer peptide, present in the α -gliadin fraction of wheat flour and recognized as the most immunogenic peptide present in wheat (Tye-Din et al., 2010b), as it contains six copies of two overlapping T-cell epitopes: three copies of the DQ2.5-glia-α1 and three copies of the DQ2.5-glia-α2. Moreover, G12 recognize other gluten immunogenic peptides resistant to intraluminal and serum proteases that are recognized by T-cells of patients with CD (Morón et al., 2008a; Comino et al., 2013; Real et al., 2014). G12 immunodepletion experiments with hydrolyzed gluten showed that this antibody reacted with those with the highest immunoactivity for celiac patients (Moreno et al., 2016). Although lines E82 and H811 showed the lowest values for immunogenic gluten, there is a group of five RNAi lines with no significant differences between them for G12 moAb values.

The PBMCs proliferative response and release of pro-inflammatory cytokine INF-γ after exposure to flour was assayed in the seven RNAi lines and compared to BW208 wild type line and 33-mer peptide as positive controls, and to rice protein extracts as negative control. PBMCs, circulating immune cells mainly constituted by T lymphocytes and B lymphocytes, play a key role in the inflammatory system. As described above, the RNAi lines used in this study showed contrasting composition in the grain proteins related with CD and other gluten pathologies. Protein extracts from two lines used in this study (E82 and

D783) were previously assayed *in vitro* (Gil-Humanes et al., 2010), showing a pronounced reduction in the proliferative responses of gliadin-specific T-cell clones. However, the use of T cell clones allows identification of variation in the stimulation by specific peptides but do not quantify their contribution to the overall immune response to gluten in patients suffering from CD. PBMCs proliferation results and IFN-γ release reported in this study corroborate the overall immunogenicity reduction of these two lines.

PBMCs proliferation assay results showed that the protein extracts from the seven RNAi lines have impaired ability to activate stimulatory response of PBMCs from CD patients. Differences in protein composition also revealed differences in CD epitopes composition. Proteomic data showed that RNAi lines tested here are not completely devoid of CD stimulatory peptides as they still contain small amounts of gliadins as well as glutenin proteins. However, not all CD epitopes are equally immunogenic (Tye-Din et al., 2010b), and our PBMCs results agree with this. The α-gliadin fraction of gluten is known to contain the most immunogenic CD epitopes, including the 33-mer peptide which is highly resistant to digestion. Lines E82, I17 and H811 contain very low or no DQ2.5 and p31-43 α-gliadin epitopes, very low content of gluten as determined by G12, and they showed the lowest capacity of inducing PBMCs proliferation presenting no significant differences in comparison with the rice control. Lines H320 and I17 also contains comparable levels of gluten and very low or no DQ2.5 and p31-43 α-gliadin epitopes, and both differ in the PBMCs response and IFN-γ release. However, both lines are different in the DQ2.5 γ-gliadin epitopes composition.

Although gliadins from wheat seems to be the primary trigger of PBMCs proliferation and IFN-γ release, other components of the protein profile such as metabolic proteins called α-amylase/trypsin inhibitors (ATI) have been identified as important factors for the development of symptoms (Volta et al., 2013). Nevertheless, RNAi lines where NGP are highly increased, they have a lower SI suggesting that these proteins may have minor contribution to activate proliferative response of CD PBMCs.

Proteomic and PBMCs analysis agree that lines H811, I17, and specially E82 have the lowest immunogenic potential. Line E82 presents the lowest amount of prolamins as determined by HPLC, with a reduction of 99, 80 and 37% of γ-, α- and ω-gliadins, respectively. In addition, the reduction of gliadins in this line is not compensated with an increase of glutenin fraction but with NGP, suggesting a minor role of this proteins on the stimulatory effect on PMBCs. Particularly, symptoms of bread made from E82 were evaluated in comparison with gluten free bread in NCGS patients, showing no differences in the appearance of symptoms (Haro et al., 2018), meaning that the higher NGP content does not trigger symptoms in NCGS patients.

The RNAi lines described in this work are of high value for assessing further evaluations in relation to other wheat-related pathologies and intolerances as some of the NGP that appear upregulated play an important role in allergies (Huebener et al., 2015). In fact, the

analysis of the PCA shows that, not only the ATIs are not related to the proliferation of PBMCs, but also none of the other NGP studied. Other authors have suggested that CD is associated with a robust humoral response directed at a specific subset of unbound gluten proteins, especially serpins (Huebener et al., 2015; Vojdani and Vojdani, 2017). Wheat serpins are serine protease inhibitors, most of them acting as a suicide substrate for the inhibition of chymotrypsin-like proteases. Some wheat serpins have been reported that have a recognition region similar to prolamin sequences and whose possible prolamin protective function has been speculated (Østergaard et al., 2000). However, our results do not support a role of serpins increasing immunogenicity.

Conclusions

The downregulation of prolamin fractions by RNAi provided wheat lines differing in the protein composition and in the content of CD immunogenic epitopes. These RNAi lines showed a reduction in total gliadin content, in specific gliadin fractions (particularly α-gliadins for some lines), and an increment in the HMW, and in the NGP such as ATIs, serpins and triticins. Proliferation assay and INF-γ release revealed three wheat RNAi lines whose stimulatory response does not differ from rice protein extract used as gluten free negative control. These lines present a very low content of gluten as determined by G12 with a pronounced decrease in α-gliadin, containing very low or no DQ2.5 and p31-43 α-gliadin epitopes with differences in glutenins and different increments in the NGP fraction. The non-gluten protein seems not to play a key role in PBMCs proliferation and INF-γ release.

Our work highlights the utility of PBMCs as a highly useful tool for the validation of stimulatory capacity of wheat lines differing in the protein composition triggering immune response in CD.

Acknowledgments

This research was funded by the Spanish Ministry of Economy, Industry and competitiveness; grant number AGL2016-80566-P and the European Regional Development Fund (FEDER). Technical assistance from Ana García is acknowledged. We also thank the generous volunteer subjects who enrolled in the study.

Supplementary material

Supplementary table 1. CD Epitopes

Epitope name	Restricted DQ	Prolamin fraction	Sequence	Reference
DQ2.5_glia_α1a	DQ2.5	alpha	PFPQPQLPY	(Arentz-Hansen et al., 2000b)
DQ2.5_glia_α1b	DQ2.5	alpha	PYPQPQLPY	(Arentz–Hansen et al., 2002)
DQ2.5_glia_α2	DQ2.5	alpha	PQPQLPYPQ	(Arentz-Hansen et al., 2000b)
DQ2.5_glia_α3	DQ2.5	alpha	FRPQQPYPQ	(Vader et al., 2002)
DQ2.5_glia_γ1	DQ2.5	gamma	PQQSFPQQQ	(Sjöström et al., 1998)
DQ2.5_glia_γ2	DQ2.5	gamma	IQPQQPAQL	(Vader et al., 2002; Qiao et al., 2005)
DQ2.5_glia_γ2a	DQ2.5	gamma	FPQQPQQPF	(Stepniak et al., 2005)
DQ2.5_glia_γ2b	DQ2.5	gamma	YPQQPQQPF	(Stepniak et al., 2005)
DQ2.5_glia_γ3	DQ2.5	gamma	QQPQQPYPQ	(Arentz–Hansen et al., 2002)
DQ2.5_glia_γ4a	DQ2.5	gamma	SQPQQQFPQ	(Arentz–Hansen et al., 2002)
DQ2.5_glia_γ4b	DQ2.5	gamma	PQPQQQFPQ	(Qiao et al., 2005)
DQ2.5_glia_γ4c	DQ2.5	gamma	QQPQQPFPQ	(Arentz–Hansen et al., 2002)
DQ2.5_glia_γ4d	DQ2.5	gamma	PQPQQPFCQ	(Qiao, unpublished)
DQ2.5_glia_γ5	DQ2.5	gamma	QQPFPQQPQ	(Arentz–Hansen et al., 2002)
DQ2.5_glia_ω1	DQ2.5	omega	PFPQPQQPF	(Tye-Din et al., 2010b)
DQ2.5_glia_ω2	DQ2.5	omega	PQPQQPFPW	(Tye-Din et al., 2010b)
DQ2.5_glut_L1	DQ2.5	LMW	PFSQQQQPV	(Vader et al., 2002)
DQ2.5_glut_L2	DQ2.5	LMW	FSQQQQSPF	(Vader et al., 2002; Stepniak et al., 2005)
DQ2.5_hor_1	DQ2.5	hordeins	PFPQPQQPF	(Vader et al., 2003; Tye-Din et al., 2010b)
DQ2.5_hor_2	DQ2.5	hordeins	PQPQQPFPQ	(Vader et al., 2003)
DQ2.5_hor_3	DQ2.5	hordeins	PIPQQPQPY	(Tye-Din et al., 2010b)
DQ2.5_sec_1	DQ2.5	secalins	PFPQPQQPF	(Vader et al., 2003; Tye-Din et al., 2010b)
DQ2.5_sec_2	DQ2.5	secalins	PQPQQPFPQ	(Vader et al., 2003)
DQ2.5_ave_1a	DQ2.5	avenins	PYPEQQEPF	(Vader et al., 2003; Arentz-Hansen et al., 2004)
DQ2.5_ave_1b	DQ2.5	avenins	PYPEQQQPF	(Vader et al., 2003; Arentz-Hansen et al., 2004)
DQ2.2_glut_L1	DQ2.2	LMW	PFSQQQQPV	(Bodd et al., 2012)
DQ2.2_glia_α1	DQ2.2	alpha	QGSVQPQQL	(Bergseng et al., 2015)

DQ2.2_glia_α2	DQ2.2	alpha	QYSQPQQPI	(Bergseng et al., 2015)
DQ8_glia_α1	DQ8	alpha	QGSFQPSQQ	(van de Wal et al., 1998)
DQ8_glia_γ1a	DQ8	gamma	QQPQQPFPQ	(Tollefsen et al., 2006a)
DQ8_glia_γ1b	DQ8	gamma	QQPQQPYPQ	(Tollefsen et al., 2006a)
DQ8_glut_H1	DQ8	HMW	QGYYPTSPQ	(van de Wal et al., 1999)
DQ8.5_glia_α1	DQ8	alpha	QGSFQPSQQ	(Kooy-Winkelaar et al., 2011)
DQ8.5_glia_γ1	DQ8	gamma	PQQSFPQQQ	(Kooy-Winkelaar et al., 2011)
DQ8.5_glut_H1	DQ8	HMW	QGYYPTSPQ	(Kooy-Winkelaar et al., 2011)
p31-43L	p31-43	alpha	LGQQQPFPPQQPY	(Maiuri et al., 1996)
p31-43P	p31-43	alpha	PGQQQPFPPQQPY	(Maiuri et al., 2003)

Supplementary table 2. Clinical data of patients with celiac disease

Patient	Age	Sex	AATG (IgA)	AAEM	Atrophy grade (Marsh criteria)	HLA-DQB1
Coeliac 1	4	Female	>200	+	IV	0201-0202
Coeliac 2	4	Female	>200	+	III C	0201-0202
Coeliac 3	1	Female	>150	+	III B	0301-0302
Coeliac 4	3	Female	>50	+	III A	0201-0603
Coeliac 5	12	Male	>200	+	IV	0201-0503
Coeliac 6	7	Male	>125	+	III A	0201-0301
Coeliac 7	1	Male	>125	+	II	0201-0602
Coeliac 8	5	Female	>125	+	III A	0201-0501
Coeliac 9	10	Male	>200	+	IV	0201-0301
Coeliac 10	2	Female	>125	+	III B	0301-0302
Coeliac 11	10	Female	>125	+	III B	0201-0202
Coeliac 12	2	Female	25	+	III A	0301-0302
Coeliac 13	3	Male	>125	+	III C	0201-0604
Coeliac 14	5	Male	>200	+	III C	0201-0202
Coeliac 15	6	Female	>150	+	III A	0201-0501
Coeliac 16	8	Female	>125	+	III A	0201-0503
Coeliac 17	2	Male	>125	+	III B	0201-0301
Coeliac 18	11	Female	>200	+	IV	0201-0202
Coeliac 19	7	Male	>200	+	III C	0201-0501
Coeliac 20	9	Female	>125	+	III B	0201-0301
Coeliac 21	6	Female	>150	+	III A	0201-0301
Coeliac 22	12	Female	>150	+	III C	0201-0202
Coeliac 23	7	Male	>200	+	IV	0201-0202
Coeliac 24	10	Female	>50	+	III B	0201-0503
Coeliac 25	11	Female	>200	+	III C	0301-0302
Coeliac 26	6	Male	>125	+	II	0201-0602
Coeliac 27	9	Female	>200	+	III B	0201-0501
Coeliac 28	3	Male	>200	+	III A	0301-0302
Coeliac 29	5	Female	>125	+	III A	0201-0301
Coeliac 30	10	Male	>150	+	III B	0201-0202
Coeliac 31	13	Female	>50	+	II	0201-0603
Coeliac 32	5	Male	>150	+	III C	0301-0302
Coeliac 33	3	Female	>200	+	IV	0201-0301
Coeliac 34	4	Male	>150	+	III C	0201-0501
Coeliac 35	6	Male	>200	+	III B	0201-0202

AAEM: antiendomysial antibody; AATG: antitransglutaminase antibody, expressed as U/ml; HLA: human leukocyte antigen

Supplementary table 3. SI Mean multiple comparison

Contrast	*Sig.*	*Difference*	*+/- Limits*

Bw208 - D623	*	1,85333	1,4108
Bw208 - D783	*	2,31	1,4108
Bw208 - E82	*	4,12333	1,4108
Bw208 - H754	*	2,60667	1,4108
Bw208 - H811	*	3,11667	1,4108
Bw208 - H320	*	2,97333	1,4108
Bw208 - IL7	*	3,85667	1,4108
Bw208 - 33-mer	*	-5,25333	1,4108
Bw208 - Control	*	4,44	1,4108
D623 - D783		0,456667	1,4108
D623 - E82	*	2,27	1,4108
D623 - H754		0,753333	1,4108
D623 - H811		1,26333	1,4108
D623 - H320		1,12	1,4108
D623 - IL7	*	2,00333	1,4108
D623 - 33-mer	*	-7,10667	1,4108
D623 - Control	*	2,58667	1,4108
D783 - E82	*	1,81333	1,4108
D783 - H754		0,296667	1,4108
D783 - H811		0,806667	1,4108
D783 - H320		0,663333	1,4108
D783 - IL7	*	1,54667	1,4108
D783 - 33-mer	*	-7,56333	1,4108
D783 - Control	*	2,13	1,4108
E82 - H754	*	-1,51667	1,4108
E82 - H811		-1,00667	1,4108
E82 - H320		-1,15	1,4108
E82 - IL7		-0,266667	1,4108
E82 - 33-mer	*	-9,37667	1,4108
E82 - Control		0,316667	1,4108
H754 - H811		0,51	1,4108
H754 - H320		0,366667	1,4108
H754 - IL7		1,25	1,4108
H754 - 33-mer	*	-7,86	1,4108
H754 - Control	*	1,83333	1,4108
H811 - H320		-0,143333	1,4108
H811 - IL7		0,74	1,4108
H811 - 33-mer	*	-8,37	1,4108
H811 - Control		1,32333	1,4108
H320 - IL7		0,883333	1,4108
H320 - 33-mer	*	-8,22667	1,4108
H320 - Control	*	1,46667	1,4108
IL7 - 33-mer	*	-9,11	1,4108
IL7 - Control		0,583333	1,4108
33-mer - Control	*	9,69333	1,4108

* denotes a statistically significant difference at the 95,0% confidence level. SI, Stimulation Index.

Supplementary table 4. IFN-γ Mean multiple comparison

Contrast	*Sig.*	*Difference*	*+/- Limits*
Bw208 - D623	*	1,62333	1,56558
Bw208 - D783	*	2,3	1,56558
Bw208 - E82	*	7,45333	1,56558
Bw208 - H754		1,27333	1,56558
Bw208 - H811	*	3,53333	1,56558
Bw208 - H320	*	3,47333	1,56558
Bw208 - IL7	*	5,60333	1,56558
Bw208 - 33-mer	*	-5,87667	1,56558
Bw208 - Control	*	7,96	1,56558
D623 - D783		0,676667	1,67367
D623 - E82	*	5,83	1,67367
D623 - H754		-0,35	1,67367
D623 - H811	*	1,91	1,67367
D623 - H320	*	1,85	1,67367
D623 - IL7	*	3,98	1,67367
D623 - 33-mer	*	-7,5	1,67367
D623 - Control	*	6,33667	1,67367
D783 - E82	*	5,15333	1,67367
D783 - H754		-1,02667	1,67367
D783 - H811		1,23333	1,67367
D783 - H320		1,17333	1,67367
D783 - IL7	*	3,30333	1,67367
D783 - 33-mer	*	-8,17667	1,67367
D783 - Control	*	5,66	1,67367
E82 - H754	*	-6,18	1,67367
E82 - H811	*	-3,92	1,67367
E82 - H320	*	-3,98	1,67367
E82 - IL7	*	-1,85	1,67367
E82 - 33-mer	*	-13,33	1,67367
E82 - Control		0,506667	1,67367
H754 - H811	*	2,26	1,67367
H754 - H320	*	2,2	1,67367
H754 - IL7	*	4,33	1,67367
H754 - 33-mer	*	-7,15	1,67367
H754 - Control	*	6,68667	1,67367
H811 - H320		-0,06	1,67367
H811 - IL7	*	2,07	1,67367
H811 - 33-mer	*	-9,41	1,67367
H811 - Control	*	4,42667	1,67367
H320 - IL7	*	2,13	1,67367
H320 - 33-mer	*	-9,35	1,67367
H320 - Control	*	4,48667	1,67367
IL7 - 33-mer	*	-11,48	1,67367
IL7 - Control	*	2,35667	1,67367
33-mer - Control	*	13,8367	1,67367

* denotes a statistically significant difference at the 95,0% confidence level.

References

Altenbach, S.B., Chang, H.-C., Yu, X.B., Seabourn, B.W., Green, P.H., and Alaedini, A. (2019). Elimination of Omega-1,2 Gliadins From Bread Wheat (Triticum aestivum) Flour: Effects on Immunogenic Potential and End-Use Quality. *Frontiers in Plant Science* 10(580). doi: 10.3389/fpls.2019.00580.

Arentz-Hansen, H., Fleckenstein, B., Molberg, Ø., Scott, H., Koning, F., Jung, G., et al. (2004). The molecular basis for oat intolerance in patients with celiac disease. *PLoS medicine* 1(1), e1-e1. doi: 10.1371/journal.pmed.0010001.

Arentz-Hansen, H., Körner, R., Molberg, Ø., Quarsten, H., Vader, W., Kooy, Y.M.C., et al. (2000). The Intestinal T Cell Response to α-Gliadin in Adult Celiac Disease Is Focused on a Single Deamidated Glutamine Targeted by Tissue Transglutaminase. *The Journal of Experimental Medicine* 191(4), 603-612. doi: 10.1084/jem.191.4.603.

Arentz–Hansen, H., McAdam, S.N., Molberg, Ø., Fleckenstein, B., Lundin, K.E.A., Jørgensen, T.J.D., et al. (2002). Celiac lesion T cells recognize epitopes that cluster in regions of gliadins rich in proline residues. *Gastroenterology* 123(3), 803-809. doi: 10.1053/gast.2002.35381.

Barro, F., Iehisa, J.C.M., Giménez, M.J., García-Molina, M.D., Ozuna, C.V., Comino, I., et al. (2016). Targeting of prolamins by RNAi in bread wheat: effectiveness of seven silencing-fragment combinations for obtaining lines devoid of coeliac disease epitopes from highly immunogenic gliadins. *Plant Biotechnology Journal* 14(3), 986-996. doi: 10.1111/pbi.12455.

Bergseng, E., Dørum, S., Arntzen, M.Ø., Nielsen, M., Nygård, S., Buus, S., et al. (2015). Different binding motifs of the celiac disease-associated HLA molecules DQ2.5, DQ2.2, and DQ7.5 revealed by relative quantitative proteomics of endogenous peptide repertoires. *Immunogenetics* 67(2), 73-84. doi: 10.1007/s00251-014-0819-9.

Bernstein, E., Denli, A.M., and Hannon, G.J. (2001). The rest is silence. RNA 7(11), 1509-1521.

Bodd, M., Kim, C.Y., Lundin, K.E.A., and Sollid, L.M. (2012). T-Cell Response to Gluten in Patients With HLA-DQ2.2 Reveals Requirement of Peptide-MHC Stability in Celiac Disease. *Gastroenterology* 142(3), 552-561. doi: 10.1053/j.gastro.2011.11.021.

Brandtzaeg, P. (2006). The changing immunological paradigm in coeliac disease. *Immunology letters* 105(2), 127-139. doi: 10.1016/j.imlet.2006.03.004.

Comino, I., Real, A., de Lourdes Moreno, M., Montes, R., Cebolla, Á., and Sousa, C. (2013). Immunological determination of gliadin 33-mer equivalent peptides in beers as a specific and practical analytical method to assess safety for celiac patients. *Journal of the Science of Food and Agriculture* 93(4), 933-943. doi: 10.1002/jsfa.5830.

Choung, R.S., Larson, S.A., Khaleghi, S., Rubio-Tapia, A., Ovsyannikova, I.G., King, K.S., et al. (2017). Prevalence and Morbidity of Undiagnosed Celiac Disease From a Community-Based Study. *Gastroenterology* 152(4), 830-839.e835. doi: 10.1053/j.gastro.2016.11.043.

Dale, H.F., Biesiekierski, J.R., and Lied, G.A. (2019). Non-coeliac gluten sensitivity and the spectrum of gluten-related disorders: an updated overview. *Nutrition Research Reviews* 32(1), 28-37. doi: 10.1017/S095442241800015X.

De Giorgio, R., Volta, U., and Gibson, P.R. (2016). Sensitivity to wheat, gluten and FODMAPs in IBS: facts or fiction? *Gut* 65(1), 169. doi: 10.1136/gutjnl-2015-309757.

Fallang, L.-E., Bergseng, E., Hotta, K., Berg-Larsen, A., Kim, C.-Y., and Sollid, L.M. (2009). Differences in the risk of celiac disease associated with HLA-DQ2.5 or HLA-DQ2.2 are related to sustained gluten antigen presentation. *Nature Immunology* 10, 1096. doi: 10.1038/ni.1780.

Fasano, A., Sapone, A., Zevallos, V., and Schuppan, D. (2015). Nonceliac Gluten Sensitivity. *Gastroenterology* 148(6), 1195-1204. doi: 10.1053/j.gastro.2014.12.049.

García-Molina, M.D., Muccilli, V., Saletti, R., Foti, S., Masci, S., and Barro, F. (2017). Comparative proteomic analysis of two transgenic low-gliadin wheat lines and non-transgenic wheat control. *Journal of Proteomics* 165, 102-112. doi: 10.1016/j.jprot.2017.06.010.

Gil-Humanes, J., Pistón, F., Altamirano-Fortoul, R., Real, A., Comino, I., Sousa, C., et al. (2014a). Reduced-Gliadin Wheat Bread: An Alternative to the Gluten-Free Diet for Consumers Suffering Gluten-Related Pathologies. *PLOS ONE* 9(3), e90898. doi: 10.1371/journal.pone.0090898.

Gil-Humanes, J., Pistón, F., Hernando, A., Alvarez, J.B., Shewry, P.R., and Barro, F. (2008). Silencing of γ-gliadins by RNA interference (RNAi) in bread wheat. *Journal of Cereal Science* 48(3), 565-568. doi: 10.1016/j.jcs.2008.03.005.

Gil-Humanes, J., Pistón, F., Shewry, P.R., Tosi, P., and Barro, F. (2011). Suppression of gliadins results in altered protein body morphology in wheat. *Journal of Experimental Botany* 62(12), 4203-4213. doi: 10.1093/jxb/err119.

Gil-Humanes, J., Pistón, F., Tollefsen, S., Sollid, L.M., and Barro, F. (2010). Effective shutdown in the expression of celiac disease-related wheat gliadin T-cell epitopes by RNA interference. *PNAS* 107(39), 7023–17028. doi: 10.1073/pnas.1007773107.

Hammond, S.M., Bernstein, E., Beach, D., and Hannon, G.J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *Nature* 404(6775), 293-296. doi: 10.1038/35005107.

Hansen, L.B.S., Roager, H.M., Søndertoft, N.B., Gøbel, R.J., Kristensen, M., Vallès-Colomer, M., et al. (2018). A low-gluten diet induces changes in the intestinal microbiome of healthy Danish adults. *Nature Communications* 9(1), 4630. doi: 10.1038/s41467-018-07019-x.

Haro, C., Villatoro, M., Vaquero, L., Pastor, J., Giménez, M.J., Ozuna, C.V., et al. (2018). The Dietary Intervention of Transgenic Low-Gliadin Wheat Bread in Patients with Non-Celiac Gluten Sensitivity (NCGS) Showed No Differences with Gluten Free Diet (GFD) but Provides Better Gut Microbiota Profile. *Nutrients* 10(12), 1964. doi: 10.3390/nu10121964.

Holtmeier, W., and Caspary, W.F. (2006). Celiac disease. Orphanet J Rare Dis 1, 3. doi: 10.1186/1750-1172-1-3.

Huebener, S., Tanaka, C.K., Uhde, M., Zone, J.J., Vensel, W.H., Kasarda, D.D., et al. (2015). Specific Nongluten Proteins of Wheat Are Novel Target Antigens in Celiac Disease Humoral Response. *Journal of Proteome Research* 14(1), 503-511. doi: 10.1021/pr500809b.

Junker, Y., Zeissig, S., Kim, S.-J., Barisani, D., Wieser, H., Leffler, D.A., et al. (2012). Wheat amylase trypsin inhibitors drive intestinal inflammation via activation of toll-like receptor 4. *Journal of experimental medicine* 209(13), 2395-2408. doi: 10.1084/jem.20102660.

Kooy-Winkelaar, Y., van Lummel, M., Moustakas, A.K., Schweizer, J., Mearin, M.L., Mulder, C.J., et al. (2011). Gluten-Specific T Cells Cross-React between HLA-DQ8 and the HLA-DQ2α/DQ8β Transdimer. *The Journal of Immunology* 187(10), 5123. doi: 10.4049/jimmunol.1101179.

Larré, C., Lupi, R., Gombaud, G., Brossard, C., Branlard, G., Moneret-Vautrin, D.A., et al. (2011). Assessment of allergenicity of diploid and hexaploid wheat genotypes: Identification of allergens in the albumin/globulin fraction. *Journal of Proteomics* 74(8), 1279-1289. doi: 10.1016/j.jprot.2011.03.014.

Lebwohl, B., Sanders, D.S., and Green, P.H.R. (2018). Coeliac disease. *The Lancet* 391(10115), 70-81. doi: 10.1016/S0140-6736(17)31796-8.

Maiuri, L., Ciacci, C., Auricchio, S., Brown, V., Quaratino, S., and Londei, M. (2000). Interleukin 15 mediates epithelial changes in celiac disease. *Gastroenterology* 119(4), 996-1006. doi: 10.1053/gast.2000.18149.

Maiuri, L., Ciacci, C., Ricciardelli, I., Vacca, L., Raia, V., Auricchio, S., et al. (2003). Association between innate response to gliadin and activation of pathogenic T cells in coeliac disease. *The Lancet* 362(9377), 30-37. doi: 10.1016/S0140-6736(03)13803-2.

Maiuri, L., Picarelli, A., Boirivant, M., Coletta, S., Mazzilli, M.C., De Vincenzi, M., et al. (1996). Definition of the initial immunologic modifications upon *in vitro* gliadin challenge in the small intestine of celiac patients. *Gastroenterology* 110(5), 1368-1378. doi: 10.1053/gast.1996.v110.pm8613040.

Marsh, M.N. (1992). Gluten, major histocompatibility complex, and the small intestine: A molecular and immunobiologic approach to the spectrum of gluten sensitivity ('celiac sprue'). *Gastroenterology* 102(1), 330-354. doi: 10.1016/0016-5085(92)91819-P.

Matsuo, H., Yokooji, T., and Taogoshi, T. (2015). Common food allergens and their IgE-binding epitopes. *Allergology International* 64(4), 332-343. doi: 10.1016/j.alit.2015.06.009.

Moreno, M.d.L., Muñoz-Suano, A., López-Casado, M.Á., Torres, M.I., Sousa, C., and Cebolla, Á. (2016). Selective capture of most celiac immunogenic peptides from hydrolyzed gluten proteins. *Food Chemistry* 205, 36-42. doi: 10.1016/j.foodchem.2016.02.066.

Morón, B., Bethune, M.T., Comino, I., Manyani, H., Ferragud, M., López, M.C., et al. (2008a). Toward the Assessment of Food Toxicity for Celiac *Patients: Characterization of Monoclonal Antibodies to a Main Immunogenic Gluten Peptide. PLOS* ONE 3(5), e2294. doi: 10.1371/journal.pone.0002294.

Morón, B., Cebolla, Á., Manyani, H., Álvarez-Maqueda, M., Megías, M., Thomas, M.d.C., et al. (2008b). Sensitive detection of cereal fractions that are toxic to celiac disease patients by using monoclonal antibodies to a main immunogenic wheat peptide. *American Journal of Clinical Nutrition* 87(2), 405-414. doi: 10.1093/ajcn/87.2.405.

Oberhuber, G., Granditsch, G., and Vogelsang, H. (1999). The histopathology of coeliac disease: time for a standardized report scheme for pathologists. *European Journal of Gastroenterology & Hepatology* 11(10), 1185.

Østergaard, H., Rasmussen, S.K., Roberts, T.H., and Hejgaard, J. (2000). Inhibitory Serpins from Wheat Grain with Reactive Centers Resembling Glutamine-rich Repeats of Prolamin Storage Proteins: Cloning and characterization of five major molecular forms. *Journal of Biological Chemistry* 275(43), 33272-33279. doi: 10.1074/jbc.M004633200.

Ozuna, C.V., Iehisa, J.C.M., Giménez, M.J., Alvarez, J.B., Sousa, C., and Barro, F. (2015). Diversification of the celiac disease α-gliadin complex in wheat: a 33-mer peptide with six overlapping epitopes, evolved following polyploidization*. Plant Journal* 82(5), 794-805. doi: 10.1111/tpj.12851.

Palosuo, K., Varjonen, E., Kekki, O.-M., Klemola, T., Kalkkinen, N., Alenius, H., et al. (2001). Wheat ω-5 gliadin is a major allergen in children with immediate allergy to ingested wheat. *Journal of Allergy and Clinical Immunology* 108(4), 634-638. doi: 10.1067/mai.2001.118602.

Pistón, F., Gil-Humanes, J., and Barro, F. (2013). Integration of promoters, inverted repeat sequences and proteomic data into a model for high silencing efficiency of coeliac disease related gliadins in bread wheat. *BMC plant biology* 13, 136-136. doi: 10.1186/1471-2229-13-136.

Pistón, F., Gil-Humanes, J., Rodríguez-Quijano, M., and Barro, F. (2011). Down-regulating γ-gliadins in bread wheat leads to non-specific increases in other gluten proteins and has no major effect on dough gluten strength. *PloS one* 6(9), e24754-e24754. doi: 10.1371/journal.pone.0024754.

Pistón, F., León, E., Lazzeri, P.A., and Barro, F. (2008). Isolation of two storage protein promoters from Hordeum chilense and characterization of their expression patterns in transgenic wheat. *Euphytica* 162(3), 371-379. doi: 10.1007/s10681-007-9530-3.

Qiao, S.-W., Bergseng, E., Molberg, Ø., Jung, G., Fleckenstein, B., and Sollid, L.M. (2005). Refining the Rules of Gliadin T Cell Epitope Binding to the Disease-Associated DQ2 Molecule in Celiac Disease: Importance of Proline Spacing and Glutamine Deamidation*. The Journal of Immunology* 175(1), 254-261. doi: 10.4049/jimmunol.175.1.254.

Real, A., Comino, I., Moreno, M.d.L., López-Casado, M.Á., Lorite, P., Torres, M.I., et al. (2014). Identification and *In Vitro* Reactivity of Celiac Immunoactive Peptides in an Apparent Gluten-Free Beer. *PLOS ONE* 9(6), e100917. doi: 10.1371/journal.pone.0100917.

Sapone, A., Lammers, K.M., Casolaro, V., Cammarota, M., Giuliano, M.T., De Rosa, M., et al. (2011). Divergence of gut permeability and mucosal immune gene expression in two gluten-associated conditions: celiac disease and gluten sensitivity. *BMC Medicine* 9(1), 23. doi: 10.1186/1741-7015-9-23.

Shan, L., Molberg, Ø., Parrot, I., Hausch, F., Filiz, F., Gray, G.M., et al. (2002). Structural Basis for Gluten Intolerance in Celiac Sprue*. Science 297*(5590), 2275. doi: 10.1126/science.1074129.

Sjöström, Lundin, Molberg, Körner, Mcadam, Anthonsen, et al. (1998). Identification of a Gliadin T-Cell Epitope in Coeliac Disease: General Importance of Gliadin Deamidation for Intestinal T-Cell Recognition. *Scandinavian Journal of Immunology* 48(2), 111-115. doi: 10.1046/j.1365-3083.1998.00397.x.

Skodje, G.I., Sarna, V.K., Minelle, I.H., Rolfsen, K.L., Muir, J.G., Gibson, P.R., et al. (2018). Fructan, Rather Than Gluten, Induces Symptoms in Patients With Self-Reported Non-Celiac Gluten Sensitivity. *Gastroenterology* 154(3), 529-539.e522. doi: 10.1053/j.gastro.2017.10.040.

Sollid, L.M., Qiao, S.-W., Anderson, R.P., Gianfrani, C., and Koning, F. (2012). Nomenclature and listing of celiac disease relevant gluten T-cell epitopes restricted by HLA-DQ molecules. *Immunogenetics* 64(6), 455-460. doi: 10.1007/s00251-012-0599-z.

Stepniak, D., Vader, L.W., Kooy, Y., van Veelen, P.A., Moustakas, A., Papandreou, N.A., et al. (2005). T-cell recognition of HLA-DQ2-bound gluten peptides can be influenced by an N-terminal proline at p-1. *Immunogenetics* 57(1), 8-15. doi: 10.1007/s00251-005-0780-8.

Tjon, J.M., van Bergen, J., and Koning, F. (2010). Celiac disease: how complicated can it get? *Immunogenetics* 62(10), 641-651. doi: 10.1007/s00251-010-0465-9.

Tollefsen, S., Arentz-Hansen, H., Fleckenstein, B., Molberg, O., Ráki, M., Kwok, W.W., et al. (2006). HLA-DQ2 and -DQ8 signatures of gluten T cell epitopes in celiac disease. *The Journal of clinical investigation* 116(8), 2226-2236. doi: 10.1172/JCI27620.

Tye-Din, J.A., Stewart, J.A., Dromey, J.A., Beissbarth, T., van Heel, D.A., Tatham, A., et al. (2010b). Comprehensive, Quantitative Mapping of T Cell Epitopes in Gluten in Celiac Disease. *Science Translational Medicine* 2(41), 41ra51-41ra51. doi: 10.1126/scitranslmed.3001012.

Vader, L.W., Stepniak, D.T., Bunnik, E.M., Kooy, Y.M.C., De Haan, W., Drijfhout, J.W., et al. (2003). Characterization of cereal toxicity for celiac disease patients based on protein homology in grains. *Gastroenterology* 125(4), 1105-1113. doi: https://doi.org/10.1016/S0016-5085(03)01204-6.

Vader, W., Kooy, Y., van Veelen, P., de Ru, A., Harris, D., Benckhuijsen, W., et al. (2002). The gluten response in children with celiac disease is directed toward multiple gliadin and glutenin peptides. *Gastroenterology* 122(7), 1729-1737. doi: 10.1053/gast.2002.33606.

van de Wal, Y., Kooy, Y.M., van Veelen, P.A., Peña, S.A., Mearin, L.M., Molberg, O., et al. (1998). Small intestinal T cells of celiac disease patients recognize a natural pepsin fragment of gliadin. *Proceedings of the National Academy of Sciences of the United States of America* 95(17), 10050-10054. doi: 10.1073/pnas.95.17.10050.

van de Wal, Y., Kooy, Y.M.C., van Veelen, P., Vader, W., August, S.A., Drijfhout, J.W., et al. (1999). Glutenin is involved in the gluten-driven mucosal T cell response. *European Journal of Immunology* 29(10), 3133-3139. doi: 10.1002/(sici)1521-4141(199910)29:10<3133::aid-immu3133>3.0.co;2-g.

Vici, G., Belli, L., Biondi, M., and Polzonetti, V. (2016). Gluten free diet and nutrient deficiencies: *A review. Clinical Nutrition* 35(6), 1236-1241. doi: 10.1016/j.clnu.2016.05.002.

Vojdani, A., and Vojdani, E. (2017). Gluten and non-gluten proteins of wheat as target antigens in autism, Crohn's and celiac disease. *Journal of Cereal Science* 75, 252-260. doi: 10.1016/j.jcs.2017.04.010.

Volta, U., Caio, G., Tovoli, F., and De Giorgio, R. (2013). Non-celiac gluten sensitivity: questions still to be answered despite increasing awareness. *Cellular And Molecular Immunology* 10, 383. doi: 10.1038/cmi.2013.28.

Wen, S., Wen, N., Pang, J., Langen, G., Brew-Appiah, R.A.T., Mejias, J.H., et al. (2012). Structural genes of wheat and barley 5-methylcytosine DNA glycosylases and their potential applications for human health. *Proceedings of the National Academy of Sciences* 109(50), 20543-20548. doi: 10.1073/pnas.1217927109.

CHAPTER 4

Low-gluten, non-transgenic wheat engineered with CRISPR/Cas9

Published in:

Susana Sánchez-León, Javier Gil-Humanes, Carmen V. Ozuna, María J. Giménez, Carolina Sousa, Daniel F. Voytas, Francisco Barro. (2017) Low-gluten, non-transgenic wheat engineered with CRISPR/Cas9. Plant Biotechnology Journal, 16, 902-910. Doi:10.1111/pbi.12837

Low-gluten, non-transgenic wheat engineered with CRISPR/Cas9

Abstract

Coeliac disease is an autoimmune disorder triggered in genetically predisposed individuals by the ingestion of gluten proteins from wheat, barley and rye. The a-gliadin gene family of wheat contains four highly stimulatory peptides, of which the 33-mer is the main Immunodominant peptide in patients with coeliac. We designed two sgRNAs to target a conserved region adjacent to the coding sequence for the 33-mer in the a-gliadin genes. Twenty-one mutant lines were generated, all showing strong reduction in a-gliadins. Up to 35 different genes were mutated in one of the lines of the 45 different genes identified in the wild type, while immunoreactivity was reduced by 85%. Transgene-free lines were identified, and no off-target mutations have been detected in any of the potential targets. The low-gluten, transgene-free wheat lines described here could be used to produce low-gluten foodstuff and serve as source material to introgress this trait into elite wheat varieties.

Introduction

Wheat is one of the most widely grown crops in the world and a major component of the human diet. Wheat grain contains gluten proteins, which are responsible for the unique viscoelastic properties of wheat-derived foods; however, they also trigger certain pathologies in susceptible individuals. Amongst these, the α-gliadin family is the main protein group associated with the development of coeliac disease and non-coeliac gluten sensitivity, which affect more than 7% of the Western population (Mustalahti et al., 2010; Sapone et al., 2011). In bread wheat, α-gliadins are encoded by approximately 100 genes and pseudogenes (Ozuna et al., 2015) organized in tandem at the Gli-2 loci of chromosomes 6A, 6B and 6D. Traditional mutagenesis and plant breeding have failed to obtain low immunogenic wheat varieties for patients with coeliac. Here, we show that CRISPR/Cas9 technology can be used to precisely and efficiently reduce the amount of α-gliadins in the seed kernel, providing bread and durum wheat lines with reduced immunoreactivity for gluten-intolerant consumers.

Methods

sgRNAs design and plasmid construction

CRISPR/Cas9 reagents were cloned into the pANIC-6E destination vector (Mann et al., 2012) downstream the Ubiquitin1 promoter from maize. Two sgRNAs (sgAlpha-1: GCCACAAGAGCAAGTTCCAT and sgAlpha-2: GGTTGTGATGGAAATGGTTG) were designed to recognize conserved regions in the coding sequence of α-gliadins in hexaploid wheat. To synthesize the expression vectors pANIC-CR-Alpha1 and pANIC-CR-Alpha2, two Gateway-compatible donor vectors, one containing TaCas9 (pGdonor-TaCas9) and another containing the sgRNA (pGdonor-sgAlpha1 or pGdonor-sgAlpha2), were combined with pANIC-6E in a multisite Gateway recombination reaction (Supplementary figure 13). pGdonor-TaCas9 contained a wheat-codon optimized Cas9 sequence (TaCas9), with an N- and C-terminal nuclear localization signals (NLS) from the simian vacuolating virus 40 (SV40) and nucleoplasmin, respectively, and the OCS terminator sequence. pGdonor-sgAlpha contained the *Triticum aestivum* U6 RNA polymerase III promoter (TaU6) for expression of the sgRNA, followed by the gRNA sequence (Supplementary figure 13).

Plant material and genetic transformation

Transgenic lines were produced using immature scutella as explants for genetic transformation as described previously (Pistón et al., 2009). Two bread wheat lines, denoted BW208 and THA53, and one durum wheat line, cv Don Pedro (DP), were used as sources for scutellum isolation and *in vitro* culture. Plasmids carrying the sgRNAs were precipitated onto 0.6-μm gold particles at 0.75 pmol/mg gold. Regeneration medium was supplemented with 2 mg/L of PPT for selecting transgenic plants. Putative transgenic plants were then transferred to soil and grown to maturity in the greenhouse, and the presence of transformation vectors was confirmed by PCR (Supplementary table 1).

Polyacrylamide gel electrophoresis analysis

Between 6 and 12 mature wheat grains per line were crushed into a fine powder and used to extract sequentially the endosperm storage proteins. Gliadins and glutenins were then separated in A-PAGE and SDS-PAGE gels as described (Gil-Humanes et al., 2012).

Reversed-phase high-performance liquid chromatography (RP-HPLC)

Gliadins and glutenins were extracted and quantified by RP-HPLC following the protocol reported (Pistón et al., 2011). Ten half-seed biological replications were carried out for each transgenic line and wild type. Protein content was expressed as μg protein/mg flour. For each line, 10 half-grains were analyzed.

Gluten content determination by competitive ELISA

Gluten content was determined by competitive ELISA assays using two monoclonal antibodies; R5 and G12. Samples for R5 were analyzed at Centro Nacional de Biotecnología (CSIC, Campus of Cantoblanco, 28049-Madrid) as described elsewhere (Valdés et al., 2003). Samples for G12 were analyzed as described previously (Barro et al., 2016). Between three and five biological replications for each line were carried out.

Sodium dodecyl sulfate (SDS) sedimentation test

The SDS sedimentation volume was determined as described (Williams, 1988). Between two and four biological replications for each line were carried out.

DNA extraction and PCR conditions for Illumina amplicon sequencing

The Illumina MiSeq system was used for amplicon sequencing producing 2 × 280 paired-end reads. PCR amplification was carried out using the forward primer aGli900F1 and the reverse primer 33mer1R2_ok (Supplementary table 1) with following conditions: 94 °C for 1 min followed by 30 cycles at 94 °C for 15 s, 62 °C for 45 s and 72 °C for 1 min with final extension at 72 °C for 2 min. For PCR amplification, 5 ng of DNA in a 25 µL volume reaction with the following final concentrations: 1× FastStart buffer, 200 nm forward and reverse primers, 200 µm dNTP mix, 1.25 units of FastStart High Fidelity polymerase (Roche Diagnostics, Mannheim, Germany) was used. Preparation of the Illumina amplicon library and sequencing was carried out at the Unidad de Genómica Cantoblanco of Fundación Parque Científico de Madrid (FPCM, Spain). The range of amplicon lengths was check using the Agilent 2100 Bioanalyzer system (Agilent Technologies, Santa Clara, CA 95051).

Amplicon sequence clustering

Fifty-three samples were subjected to amplicon sequencing: six samples corresponded to wild-type DNA (2 BW208, 2 DP and 2 THA53), and 47 to DNA from transgenic lines (Supplementary table 2 and 4). In total, 33.817 millions of reads were obtained. For clustering, the USEARCH software v8.0.1517 (Edgar, 2010) was used. Merging of paired-end reads was using the -fastq_mergepairs command, and for quality filtering by expected errors -fastq_filter (-fastq_maxee 1) commands were used (Edgar and Flyvbjerg, 2015). Then, all 11.676 millions of cleaned and filtered reads were clustered with -cluster_otus mode, 100% homology and -search_exact command for mapping and to extract the consensus sequence for each cluster. To extract the high-confidence amplicon variants for each sample, samples with less than five reads in a given cluster were removed from that cluster. As clustering was at 100%, consensus clusters were considered as unique genes. As

samples have different numbers of reads, frequencies were calculated for each sample by dividing the number of reads for a given amplicon gene (n) by the total count for a sample (N). Then, all gene sequences were processed using Geneious version 9.1.4 (Biomatters Ltd., Auckland, New Zealand; available at http://www.gene-ious.com/). First, a reference unique gene library was constructed for each of the wild-type lines. Genes with different lengths were used as reference sequences for aligning and mapping of amplicon genes from mutant lines. Second, genes present in mutant lines were aligned and mapped to reference gene library constructed previously using the BBmap aligner (https://sourceforge.net/projects/bbmap/). MAFFT software v7.222 (Katoh et al., 2002) and Fast Tree software (Price et al., 2010) were used for multiple sequence alignment and maximum-likelihood phylogenetic trees, respectively, to determine the corresponding non mutated sequences.

PCR amplification of γ- and ω-gliadin genes and sequencing by Sanger

The gene-specific primers for Sanger sequencing of γ- and ω-gliadin genes are in Supplementary table 1. These primers amplified from signal peptide in the 5′ to end of the coding region in the 3′. The complete γ- and ω-gliadin genes were amplified by PCR as follow: 94 °C for 4 min followed by 35 cycles at 94 °C for 15 s, 60 °C or 66 °C (γ-gliadins and ω-gliadins, respectively) for 1 min and 72 °C for 1 min 30 s, with final extension at 72 °C for 7 min. For PCR amplification, 200 ng of DNA in a 25 µL volume reaction consisting of 400 nm forward and reverse primers, 320 µm dNTP mix, a mixture of 0.013 units Pfu DNA polymerase (Biotools, B&M Labs, Madrid, Spain) and 0.650 units Taq DNA polymerase (Biotools) was used. PCR products were checked by 1% agarose gel electrophoresis.

Full-length DNA sequences were ligated into pGEM-T Easy vector (Promega, Madison, WI) and cloned into Escherichia coli DH5α cells. Sequencing was carried out by Stab Vida (Caparica, Portugal). We sequenced 102 clones (35 wild types and 67 mutant lines) and 78 clones (26 wild types and 52 mutant lines) for the γ- and ω-gliadin genes, respectively.

Detection of bar and Cas9 genes, and other DNA plasmid regions

PCR was performed to detect insertions of plasmid DNA using primers listed in Supplementary table 1. For detection of bar and Cas9 genes, and PVS1 stability (sta) region, Octopine synthase polyA signal, kanamycin resistance gene and *Panicum virgatum* ubiquitin 1 promoter, 300 ng of DNA was used in a 25 µL volume reaction, consisting of 400 nm forward and reverse primers, 320 µm dNTP mix and 0.650 units Taq DNA polymerase (Biotools, Madrid, Spain). PCR conditions were as follows: 94 °C for 4 min followed by 35 cycles at 94 °C for 15 s, 58 °C for 45 s or 30 s for Cas9 and the other genes, respectively, and 72 °C for 1 min 30 s with final extension at 72 °C for 7 min.

Specifics primers (Supplementary table 1) were designed to be used with aGli900 Forward primer for the amplification of each insertion detected by deep sequencing. PCR conditions

for the amplification of insertions were as followed: 94 °C for 5 min followed by 35 cycles at 94 °C for 30 s, 60 °C for 30 s and 72 °C for 30 s with a final extension at 72 °C for 5 min. For PCR amplification, 300 ng of DNA in a 25 µL volume reaction consisting of 400 nm forward and reverse primers, 320 µm dNTP mix and 0.650 units Taq DNA polymerase (Biotools) was used. All PCR products were checked by 1% agarose gel electrophoresis.

Analysis of off-target mutations

Potential off-targets in the wheat genome were detected by two different methods. First, we performed an in silico search of the minimal active sequence of the sgRNAs in the prolamin genes (except α-gliadins) deposited in the GenBank database (http://www.ncbi.nlm.nih.gov/). We used the seed sequence (12 nt upstream the PAM sequence) of the sgAlpha-1 and sgAlpha-2 plus the NGG PAM sequence and searched for homology in 179 γ-gliadins, 15 ω-gliadins, 40 HMW glutenins and 239 LMW glutenins, allowing up to two mismatches in the seed sequence. Then, we expanded our in silico search for off-target sites to the whole genome of wheat by searching for perfect matches of the seed sequence (12 nt) plus PAM in the reference genome of bread wheat (http://plants.ensembl.org/index.html).

Potential off-targeted genes were characterized in 3 T1 mutant plants (T544, T545 and T553). Specific primers were designed to PCR amplify a 267 to 323 bp fragment encompassing the potential off-target site of sgAlpha-1 or sgAlpha-2 in the identified genes (Traes_7BL_F621D9B9E (MADS box transcription factor), Traes_2AS_D659E88E9.1, Traes_2AS_8FCC59363.1 and the gene family of LMW glutenins) (Supplementary table 1). Amplicons were cloned, and between 24 and 39 clones were sequenced for each of the genes.

Matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) analysis

Gliadin fractions were extracted from wheat flours using 60% ethanol for 1 h at room temperature in a rotary shaker and centrifuged at 12000 g for 5 min at room temperature. Previously, samples were washed twice with 0.5 m NaCl for 30 min at 4 °C in a rotary shaker and centrifuged at 12000 g for 5 min at 4 °C, to remove albumins/globulins fraction.

The ethanolic supernatants obtained for each sample were diluted at 1 : 1 ratio (v/v) with matrix solution (10 mg/mL 2,5 Dihydroxyacetophenone in 50% aqueous acetonitrile and 100 mM ammonium citrate). A 1.0 µL aliquot of this mixture was manually deposited onto a 386-well OptiTOF™ Plate (Sciex) and allowed to dry at room temperature. For MALDI-TOF/TOF analysis, samples were automatically acquired in an ABi 4800 MALDI-TOF/TOF mass spectrometer (Sciex) in positive ion linear mode (the ion acceleration voltage was 25 kV for MS acquisition). The detection mass range was set between 1500 and 80 000 m/z.

Statistical analysis

Data were analyzed with the statistical software Statistix v10 (Analytical software, PO Box 12185, Tallahassee, FL 32317). The differences in the data were assessed using analysis of the variance (ANOVA), followed by the two-tailed Dunnett's post hoc test for median multiple comparisons. P values lower than 0.05 were considered significant. Shapiro–Wilk normality test was used to verify that data was normally distributed, and logarithmic or Box Cox transformations were applied whenever a variable did not pass the test. Figures were drawn using the Microsoft Excel and PowerPoint software (Microsoft Corporation).

Results and discussion

To precisely modify the immunoreactive α-gliadin genes, we designed two sgRNAs (sgAlpha-1 and sgAlpha-2) (Figure 1a) to target conserved regions adjacent to the coding sequence for the immunodominant epitope in wheat gluten, a protease-resistant, 33-amino acid peptide that contains six overlapping copies of three distinct, tandemly organized epitopes (DQ2.5-glia-α1a, PFPQPELPY; DQ2.5-glia-α2, PQPELPYPQ; and DQ2.5-glia-α1b, PYPQPELPY) (Tye-Din et al., 2010b). The CRISPR/Cas9 constructs were transformed into two bread wheat (BW028 and TAH53) and one durum wheat (DP) cultivars, resulting in twenty-one (15 bread wheat and 6 durum wheat) T0 transgenic lines. DNA was isolated from leaves of 17 T1 transgenic plants (5 BW208, 4 TAH53, and 8 DP) and the corresponding wild-type varieties, and PCR amplicons encompassing the sgAlpha-1 and sgAlpha-2 target sites were subjected to Illumina high-throughput DNA sequencing (Figure 1a, Supplementary table 1). We observed considerable variability in the bread wheat and durum wheat wild-type sequences, due to randomly distributed SNPs and differences in the number of encoded epitopes in the 33-mer region (Supplementary figure 1). As expected, a number of sequences were pseudogenes with premature stop codons, and frameshift mutations in the C-terminus (Supplementary figure 2). We found 45, 52 and 43 different α-gliadin sequences that were highly represented (frequencies higher than 0.3%) in BW208, THA53 and DP, respectively. Of these, 35, 13 and 29 were, respectively, mutated by CRISPR/Cas9 (Supplementary figure 3). The mutation spectrum in the α-gliadins was characterized in the various T1 transgenic plants (Figure 1b–d, Supplementary table 2, Supplementary figures 4–6). Due to the presence of the Cas9 expression vector in some of the mutant lines, the frequency of mutations observed might be overestimated, as a consequence of somatic mutations. However, in most cases we observed similar mutation frequencies in T1 plants generated from the same T0 plant, with or without Cas9—that is V467 (+Cas9, 5.18% NHEJ) and V468 (−Cas9, 5.17% NHEJ), both derived from T0 plant #20 (Supplementary table 2).

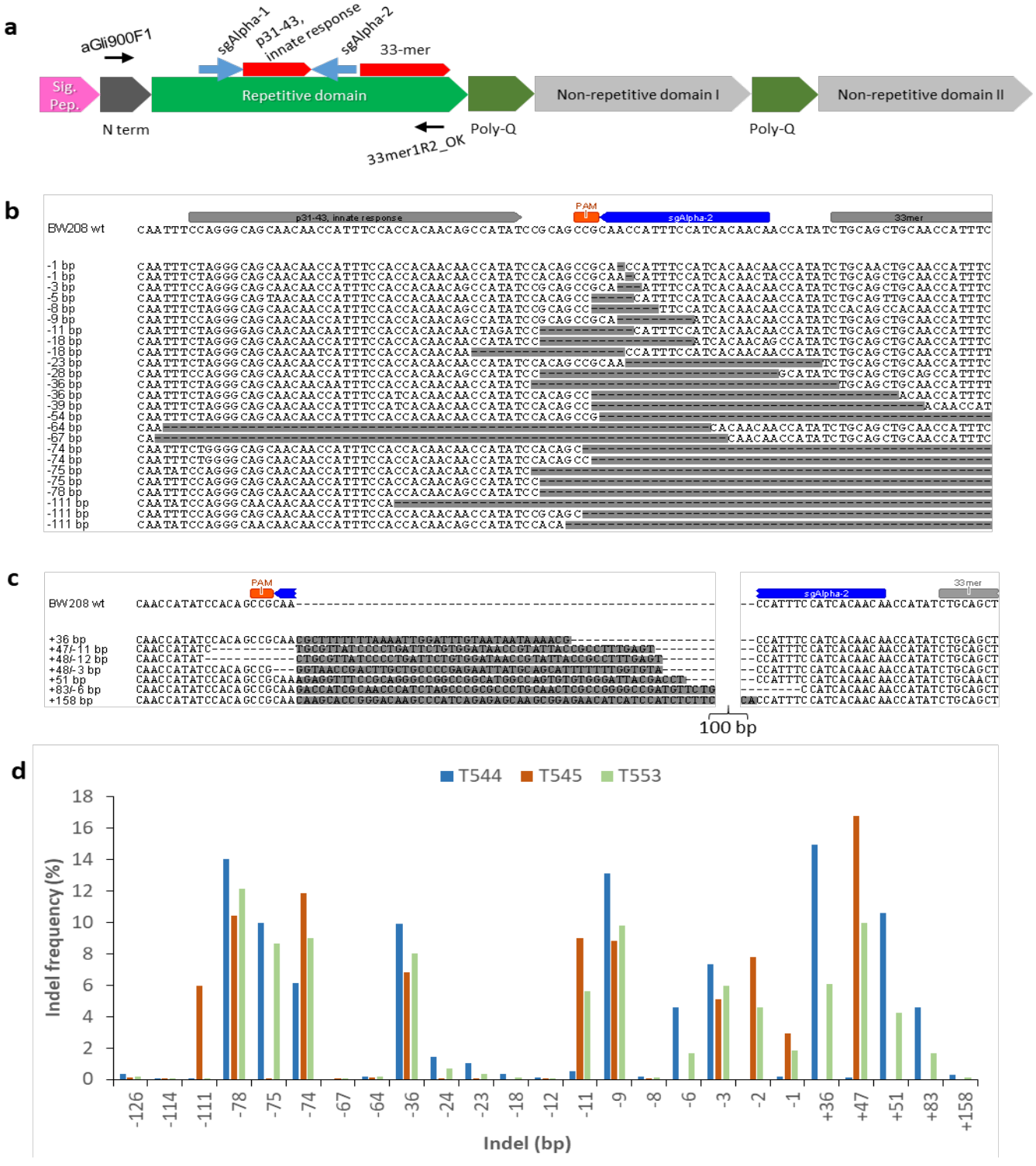


Figure 1. Gene editing of α-gliadins in bread wheat. (a) Schematic of a typical α-gliadin gene indicating the different protein domains. Two of the peptide sequences involved in gluten intolerance (p31-43 and the 33-mer) are represented by red arrows, whereas the target sequences for the sgRNAs (sgAlpha-1 and sgAlpha-2) are represented by blue arrows. Black arrows indicate primers used for Illumina sequencing. (b–d) Illumina sequencing of the α-gliadin genes of 3 T1 BW208 mutant lines (T544, T545 and T553) transformed with sgAlpha-2. (b) Alignment of the different deletion types found at the target locus of sgAlpha-2; (c) Alignment of the different insertions at the target locus of sgAlpha-2; and (d) frequency of the different type of insertions and deletions.

In general, sgAlpha-2 was more effective than sgAlpha-1. It should be noted that lower regeneration was observed in transgenic plants containing sgAlpha1 (0.3% transformation frequency) than plants with sgAlpha2 (1% transformation frequency). Therefore, a possible toxic effect might be also affecting the NHEJ activity of sgAlpha-1, favoring the regeneration of those lines in which the level of expression of sgAlpha1 is lower, and consequently the mutation frequency. The highest mutation frequencies (62.3%–75.1%) were observed in the BW208-derived lines transformed with sgAlpha-2 (Supplementary table 2). Three of these T1 lines (T544, T545 and T553) had insertions and deletions (indels) at the target site of between +36 and +158 bp and −1 and −126 bp, respectively (Figure 1b–d). Line T545 had the highest mutation frequency of all analyzed lines: ∼75% of the sequence reads had indels (Supplementary table 2). Transgenic lines of cv DP and cv THA53 showed lower indel frequencies, ranging between 1.50% and 14.77% and 5.16%–7.86%, respectively. Interestingly, the typical −1bp deletion normally observed with CRISPR/Cas9 was very frequent in two of the DP sgAlpha-2 lines (22.4%–35.6%), but only represented 0.18%–2.9% of the mutations found in the BW208 sgAlpha-2 lines (Figures 1 and Supplementary figure 4). The −1bp deletion was not found in the THA53 lines or in the sgAlpha-1 lines. The +1 bp insertion, also reported as a typical mutation of CRISPR/Cas9, was only found at low frequency in one of the sgAlpha-1 lines. A possible explanation for this observation is the preference of certain types of deletions due to microhomology-mediated repair. The target sites of sgAlpha-1 and sgAlpha-2 are highly repetitive, and as shown in Supplementary figure 7, repeats between 3 and 36 bp are commonly found flanking the targeted break. The repeats could explain the bias in favor of some of the most frequent mutations observed, such as the −75 and −11 bp deletions in BW208 sgAlpha-2 lines, the −15 bp deletion in DP sgAlpha-2 lines and the −36 bp deletion in all BW208, TAH53 and DP lines. DNA insertions represented up to 19% of the total indels (Line T544, Supplementary table 2), and they were found to be either fragments of the transformation vectors or other α-gliadin genes, probably inserted by microhomology-mediated repair. These results demonstrate that high mutation frequency and specificity can be achieved using CRISPR/Cas9 to modify complex genomic loci such as the α-gliadin gene family in bread and durum wheat.

To assess the impact of the observed mutations on seed protein composition, gliadin and glutenin content in T1 half-seeds was qualitatively assessed by A-PAGE and SDS-PAGE, respectively (Figures 2a,c and Supplementary figure 8). In A-PAGE and SDS-PAGE gels, some tracks are not contiguous and have been combined from different, or the same gels (white lines) for presentation clarity. A-PAGE demonstrated that α-gliadins were strongly reduced in some of the bread and durum wheat T1 lines (e.g. plants # 6, 10, 12, 15, 17, 32 and 50), and partially reduced in others (e.g. plants # 14, 21, 28 and 48). The γ- and ω-gliadins were also strongly decreased in some lines (e.g. plants # 6, 10, 12, 15 and 21). Additional, novel bands, especially in the region of the gel corresponding to the α-gliadins, were clearly visible in the A-PAGE gels, perhaps due to truncation of α-gliadin coding sequences by mutation (Supplementary figure 8). In addition, an accumulation of one mid-

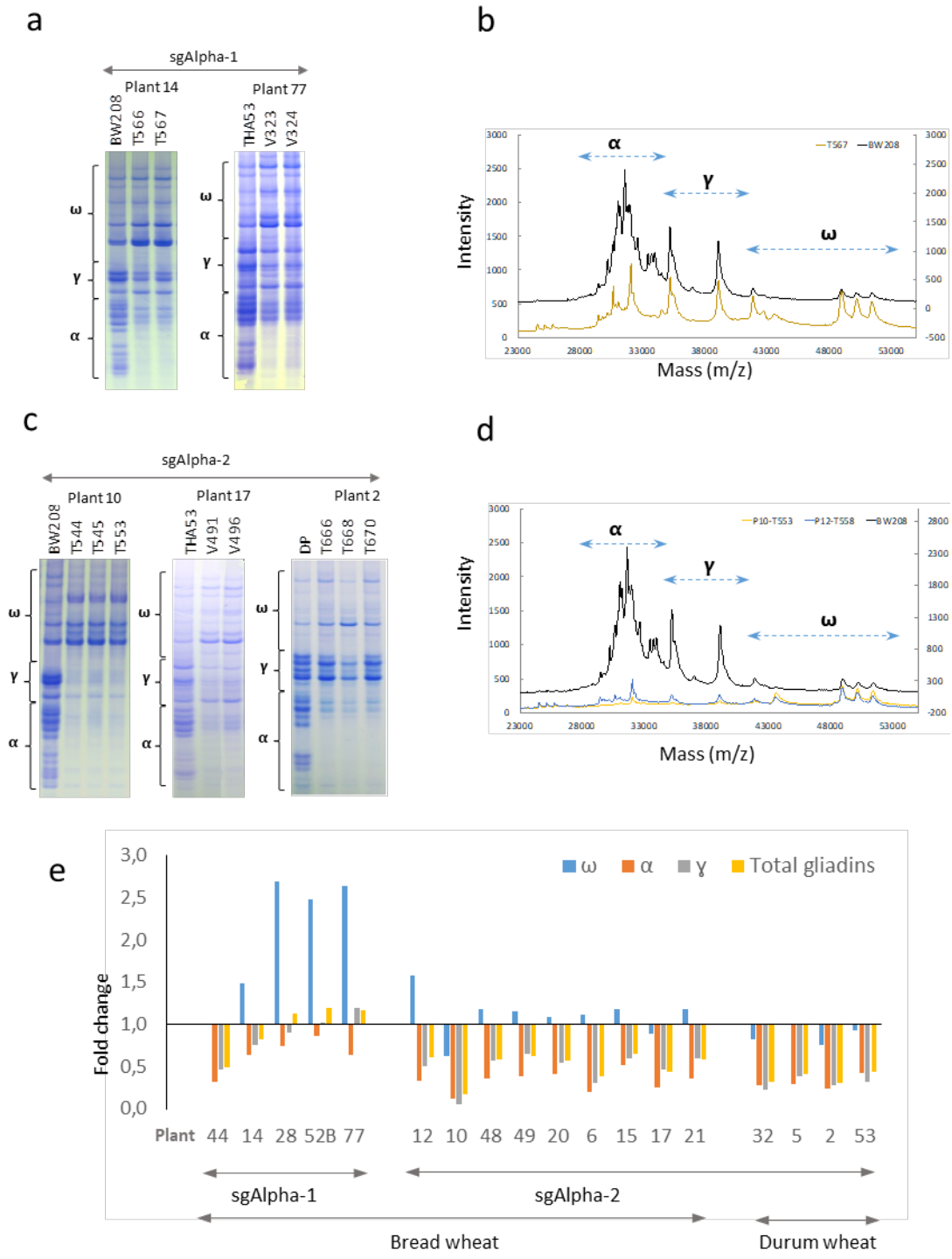


Figure 2. Characterization of sgAlpha-1 and sgAlpha-2 mutant plants. (a) A-PAGE of gliadins from sg Alpha-1 T1 half-seeds (named as T566 and T567 lines) derived from T0 plant 14, and V323 and V343 (from T0 plant 77) and the corresponding wild-type lines BW208 and THA53. Migration of α-, γ-, ω-gliadin protein bands are outlined by brackets (b) MALDI-TOF analysis of the same gliadin extract in (a) from T567 track and the BW208 wild type. Values are in absolute intensity. Left axis corresponds to T567 and the right axis to the BW208 line. The corresponding range of masses (m/z) for α-, γ-, ω-gliadins are indicated by arrows. (c) A-PAGE of gliadins from sgAlpha-2 T1 half-seeds (named as T544, T545 and T553 lines) from T0 plant 10, V491 and V496 (plant 17), T666, T668 and T670 (plant 2) and the wild-type lines BW208, THA53 and DP. (d) MALDI-TOF analysis of the same gliadin extracts in (a) from T553 track (plant 10) and T558 (plant 12, 8), and the BW208 wild type. (e) Bar graph of fold change of α-,

γ-, ω- and total gliadin fractions in bread and durum wheat transformed with sgAlpha-1 and sgAlpha-2. Values for each plant were normalized by values of the corresponding wild-type lines. Note that A-PAGE analysis is not a quantitative test, and intensity differences observed in the gels might be explained in part by differences in the amount of protein loaded and/or by differences in the staining/distaining process.

range ω-gliadin was observed in Plant 10 (Figures 2a and Supplementary figure 8). Mass spectrometry (MALDI-TOF) confirmed the sharp reduction of α-gliadins in both sgAlpha-1 and sgAlpha-2 lines, with the sgAlpha-2 lines showing a greater reduction in the number of visible peaks (Figure 2b,d). As suggested by the Illumina sequencing results, sgAlpha-2 more effective reduced the α-gliadin content, particularly in the BW208 bread wheat lines. The glutenin profile for all lines was comparable to that of the wild type; however, differences in the intensity of the two glutenin fractions were observed, suggesting a mechanism for compensatory reduction in the abundance of these proteins in response to the reduction of α-gliadins (Supplementary figure 8).

Encouraged by these results, HPLC analysis was performed to accurately quantify and characterize the different groups of gliadins and glutenins (Figure 2, Supplementary table 3). As expected, α-gliadin content was significantly reduced in most of the transgenic lines compared to the wild type (32%–82% reduction), especially in the bread and durum wheat lines transformed with sgAlpha-2. The γ-gliadins were also significantly reduced by 25%–94% in 15 of the 18 T1 lines analyzed, whereas the ω-gliadins showed the greatest variability: ω-gliadins were not affected in all four durum wheat lines, significantly up-regulated (twofold–threefold) in all four bread wheat sgAlpha-1 T1 lines (Figure 2e), and down-regulated by 33% in bread wheat Plant 10. Interestingly, this line had the highest reduction in α-gliadins (82%) and γ-gliadins (92%) and consequently showed the highest overall gliadin reduction (82%). Amongst the durum wheat lines, Plant 2 had the highest overall gliadin reduction (69%). The reduction in the gliadin content promoted a compensatory effect in glutenins, increasing the HMW fraction, especially in the BW208 and THA53 bread wheat lines (Supplementary table 3). The LMW fraction was significantly reduced only in Plant 10 and 32. Similar compensatory effects were observed previously (Gil-Humanes et al., 2011) in wheat lines in which the α-, γ- and ω-gliadins were down-regulated by RNAi. In those RNAi lines, compensatory effects provided wheat lines with no difference in the total protein content; however, changes in seed protein expression had important implications on the properties of the flour (Gil-Humanes et al., 2014b), as higher glutenin contents, particularly HMWs, are usually associated with stronger flours. The lines reported here show reduced total gliadin content (specifically the α-gliadins containing the 33-mer epitope), increased HMW glutenins and lower gli/glu ratios than the wild type.

To confirm that the altered gliadin content effectively reduced the immune reactivity of the flour, we analyzed the T2 seeds of the mutant lines with the monoclonal antibodies (moAb) R5 and G12 (Figure 3). R5 is the moAb of choice in the food industry to quantify gluten content and detects a conserved domain (QQPFP) found in most gliadins (not only the ones

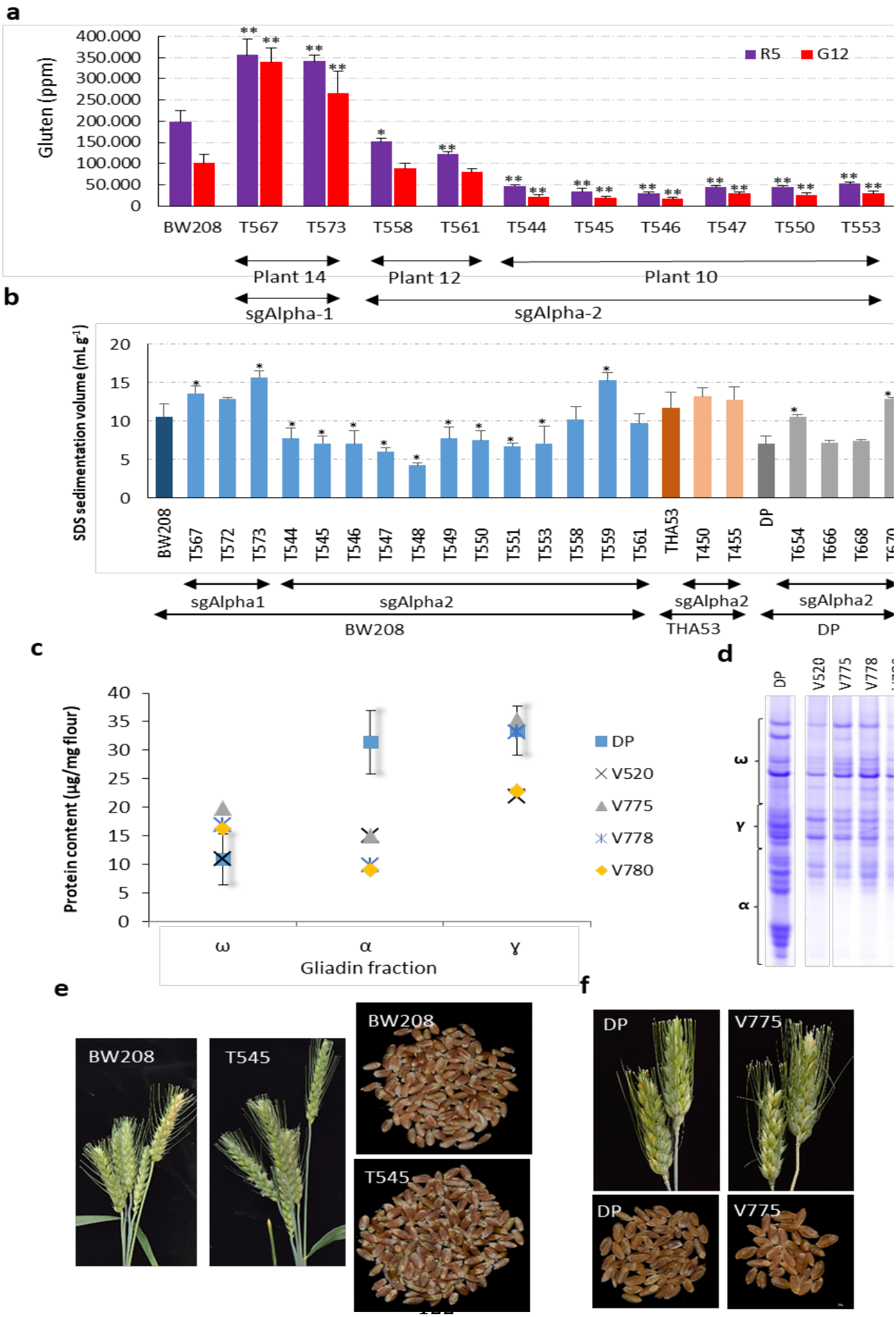

a
Gluten (ppm)
400.000
350.000
300.000
250.000
200.000
150.000
100.000
50.000
0
R5
G12
BW208
T567
T573
T558
T561
T544
T545
T546
T547
T550
T553
Plant 14
Plant 12
Plant 10
b
sgAlpha-1
sgAlpha-2
SDS sedimentation volume (mL g-1)
20
15
10
5
0
BW208
T567
T572
T573
T544
T545
T546
T547
T548
T549
T550
T551
T553
T558
T559
T561
THA53
T450
T455
DP
T654
T666
T668
sgAlpha1
sgAlpha2
sgAlpha2
sgAlpha2
BW208
THA53
DP
c
Protein content (µg/mg flour)
40
35
30
25
20
15
10
5
0
DP
V520
V775
V778
V780
ω
α
ɣ
Gliadin fraction
d
DP
V520
V775
V778
ω
γ
α
e
BW208
T545
BW208
T545
f
DP
V775
DP
V775

Figure 3. Analysis of Immune reactivity, SDS sedimentation volumes and gliadin profile of non transgenic DP-derived lines, and phenotype of sgAlpha-derived lines. (a) Analysis of T2 seeds of the sgAlpha-1 and sgAlpha-2 mutant lines with the monoclonal antibodies (moAb) R5 and G12. Error bars, mean ± SD. Statistically significant differences between each mutant line and the wild type were denoted *P < 0.05, **P < 0.01 (Tukey HSD all-pairwise comparisons test) (b) Sodium dodecyl sulfate (SDS) sedimentation test expressed as mLg-1. T2 and T3 seeds from each line were bulked and three independent biological replications analyzed. Error bars, mean ± SD. * Means are significantly different to wild types as determined by Dunnett's multiple comparisons at P < 0.05. (c) Content of the ω-, α- and γ-gliadin fraction of the non transgenic DP-derived lines. Error bars, 5% Confidence Interval of the mean value of the wild-type DP line. (d) A-PAGE of gliadins from half-seeds of the non transgenic DP-derived lines. Migration of α-, γ-, ω-gliadin protein bands are outlined by brackets. (e) Spikes and seeds of sgAlpha-2 BW208 mutant line in comparison with its wild type. (f) Spikes and seeds of sgAlpha-2 DP mutant line in comparison with its wild type.

that are immune reactive) (Valdés et al., 2003). The G12 mAb is more specific for detecting reactive epitopes, as it was developed against the 33-mer peptide (Morón et al., 2008). ELISA tests with both moAbs showed a strong reduction in gluten content in the sgAlpha-2-derived lines compared to that of the BW208 wild type. In those lines, we observed up to 85% reduction in gluten content (Line T546), and an average reduction of 66.7% and 61.7%, respectively, with the R5 and G12 moAbs. However, both moAbs revealed an increase in the gluten content for lines with sgAlpha-1. These two lines have a higher ω-gliadin content—a consequence of the knock-down of the α-gliadins (Supplementary table 3 and 4)—which could explain the observed increment in gluten content. Similar increases in gluten content when only the γ-gliadins were down-regulated by RNAi were previously reported (Gil-Humanes et al., 2008). In total, these results demonstrate that gluten immunoreactivity can be significantly reduced by editing the α-gliadin genes containing the immunodominant 33-mer epitope.

Once we had demonstrated the high efficiency of CRISPR/Cas9 to simultaneously mutate most of the α-gliadin genes, we next asked whether off-target mutations were occurring at other sites due to sgAlpha-1- and sgAlpha-2-mediated cleavage. First, we looked for possible off-target mutations in the γ- and ω-gliadin genes, as these proteins were reduced in the mutant lines. Sanger sequencing of fifty-seven clones containing γ-gliadin genes (Supplementary figure 9) and forty-three clones with ω-gliadins genes ─ twenty-four ω1,2-gliadins (Supplementary figure 10) and nineteen ω5-gliadins (Supplementary figure 11)—showed no off-target mutations. These results were confirmed by in silico search of the sgAlpha-1 and sgAlpha-2 sequences (NGG PAM plus 12 nt seed sequence, allowing for up to 2 mismatches) in the wheat prolamin genes annotated in the GenBank (Supplementary figure 12a). Additional sequencing of 11–16 clones of amplified LMW from 3 T1 mutant lines (T544, T545 and T553) showed no mutation in the only potential target site identified (Supplementary figure 12b). We therefore concluded that the observed decrease in the γ- and ω-gliadins and the glutenins in the mutant lines was not a consequence of off-target mutations. Rather, we speculate that antisense α-gliadin sequences could be expressed, resulting in the observed broad reduction in α-gliadin, as well as the down-regulation of the other gliadin proteins. Antisense sequences of α-gliadins could originate in the mutant lines as a consequence of cleavage at two target sites and inversion of the intervening DNA

sequence. We tried to detect such hypothetical inversions by predicting the inversion product and performing PCR assays (data not shown); however, none were detected in any of the tested lines.

Next, we expanded our search for off-target sites to the entire bread wheat genome (Supplementary figure 12c). Amongst all potential off-target sites (41 for sgAlpha1 and 50 for sgAlpha2), only four were annotated genes: a putative MADS box transcription factor (Traes_7BL_F621D9B9E), two genes with unknown function (Traes_2AS_D659E88E9.1, Traes_2AS_8FCC59363.1) and one gene with homology with α-gliadins (Traes_4AL_4FF5B8837). No mutations were identified in any of these genes in approximately 10 clones sequenced from each gene in the T1 mutant lines T544 (Supplementary figure 12d), T545 and T553. Collectively, these results demonstrate the high specificity of the sgRNAs designed to target the α-gliadins. Further characterization of potential off-target sites in other non-annotated genes in the genome would be necessary to confirm the lack of undesired mutations.

We next examined whether the mutations were transmitted to the next (T2) generation. Illumina sequencing of 29 T2 plants, with and without Cas9, showed heritability of the mutations (Supplementary table 4). Confirming our observations in the T1 generation, the presence/absence of the Cas9 expression vector in T2 plants did not affect the mutation frequencies (Supplementary table 4), and we believe that the variability observed between different lines can be explained by (i) stable and somatic mutagenesis due to activity of Cas9 and (ii) segregation of heterozygous stable mutations produced in the previous generations. T2 lines derived from T0 plant #10 were selected for sequencing because they had the least amount of integrated DNA at the cut sites.

The phenotype observed in the prolamin (gliadins and glutenins) content was also inherited, as assessed by evaluating 25 different T2 lines (Supplementary table 5), and 16 T3 lines (Supplementary table 6) by RP-HPLC. This demonstrated that the low-gluten trait is stable and heritable, and will enable its introgression into elite wheat varieties. As observed in T1 seeds, the HMW glutenins were also increased in T2 and T3 seeds of mutant lines (Supplementary tables 5 and 6). The HMW fraction is a major determinant of the functionality of wheat flour. We assessed the bread-making quality of the mutant lines using the SDS sedimentation test by bulking T2 and T3 seeds from each line (Figure 3b). Although some mutant lines showed higher SDS values (higher quality) than the wild-type control, we observed significant reductions in the SDS values (lower quality) in the mutant lines with the greatest reduction of gliadins. However, in most cases, SDS values in the sgAlpha2 lines were comparable to those of some RNAi lines previously reported showing 97% reduction in the gluten content (Gil-Humanes et al., 2010). Flour from those low-gluten RNAi lines showed increased stability and better tolerance to over-mixing (Gil-Humanes et al., 2014b) and allowed the production of bread with baking and sensory properties comparable to those of normal wheat flour (Gil-Humanes et al., 2014a). Consequently, one

might expect that mutant lines reported here will produce flour of a good quality and bread-making performance.

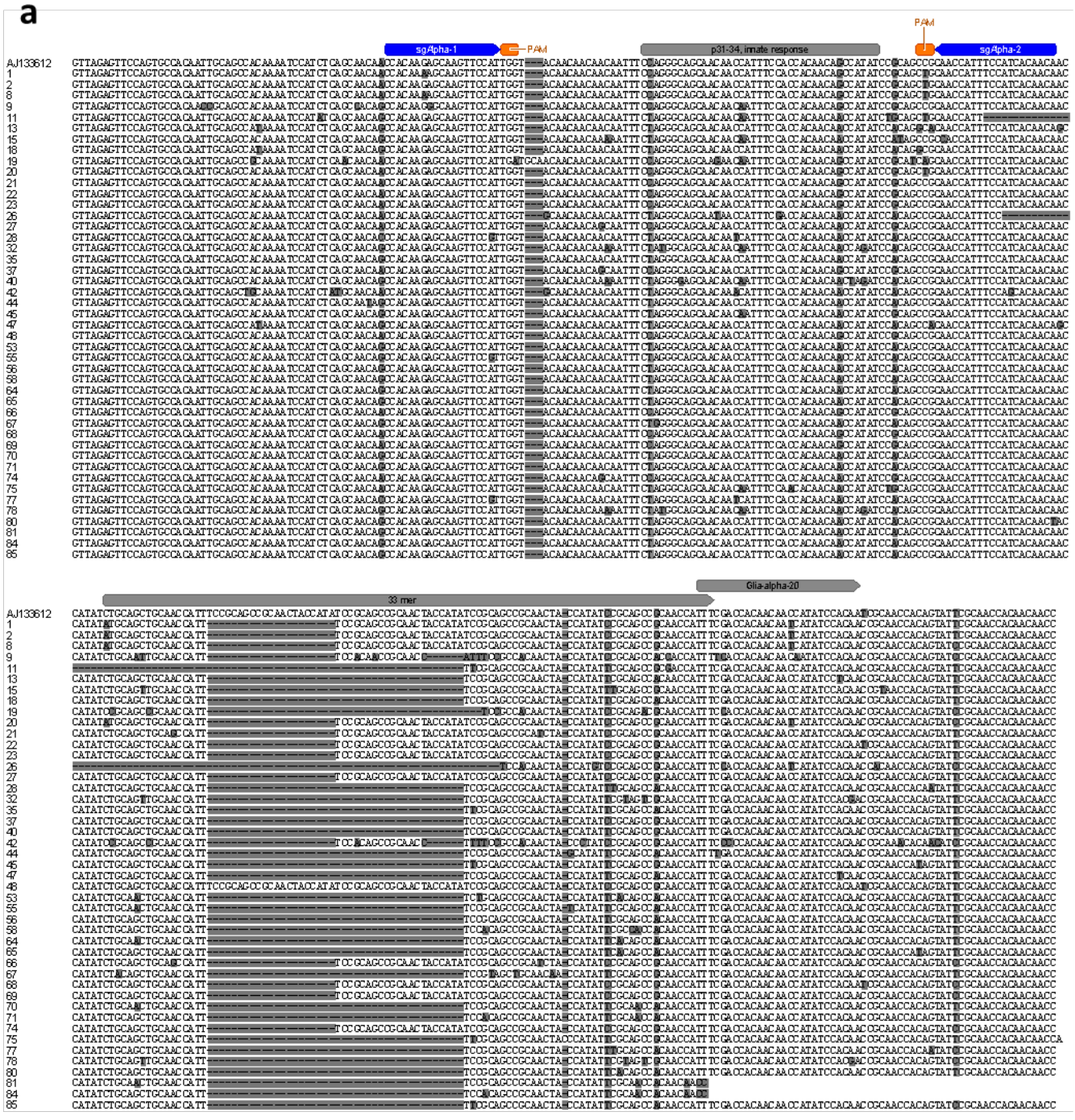
Finally, we tested whether any of the low-gluten wheat lines we generated were transgene- and insertion-free (i.e. lacked insertions at the cleavage site). We screened the T1 and T2 wheat lines by PCR and Illumina high-throughput sequencing for the presence of plasmid DNA (Supplementary figure 14). Three bread wheat (BW208) and six durum wheat (DP) T2 plants were identified as transgene-free and insertion-free (Supplementary figure 14). These non transgenic lines showed reduction in α-gliadins (Figures 3c, d, and Supplementary figure 15) showing. In all cases, all T0, T1 and T2 generations of sgAlpha-1 and sgAlpha-2 mutant bread and durum wheat were fully fertile and set seeds, and had normal chromosome numbers (Figure 3c).

We modified the coeliac disease-causing α-gliadin gene array using CRISPR/Cas9 technology to obtain non transgenic, low-gluten wheat lines. Because of the complexity of the Gli-2 locus and the high copy number of the α-gliadin genes, traditional plant breeding and mutagenesis have failed to achieve low-gluten wheat. However, CRISPR/Cas9 efficiently and precisely targeted conserved regions of the α-gliadin genes in bread and durum wheat, leading to high-frequency mutagenesis in most gene copies. Immunoreactivity of the CRISPR-edited wheat lines was reduced by 85%, as revealed the R5 and G12 ELISA tests. We previously reported the down-regulation of gliadin genes by RNAi (Gil-Humanes et al., 2010; Barro et al., 2016). Both CRIPSR/Cas9 and RNAi are highly effective for obtaining wheat lines lacking coeliac disease epitopes. However, the main advantages of the CRISPR knockouts vs RNAi are that (i) CRISPR knockouts are stable and heritable mutations that do not involve the expression of a transgene, and (ii) therefore, they provide a phenotype that is independent of environmental conditions. In addition, CRISPR/cas9 would allow different strategies to that reported in this work, that is cutting larger chromosome fragments containing gliadin genes, or even more, replacement of highly immunogenic fragments with others less toxic, keeping the gliadins functionality. In contrast, to obtain all gliadin genes mutated by CRISPR/Cas9, subsequent rounds of mutagenizing will be needed using specific sgRNAs to target the remaining gliadin genes. The low-gluten, transgene-free wheat lines described here constitute an unprecedented advance, and the resultant lines provide excellent source material for plant breeding programs to introgress the low-gluten trait into elite wheat varieties.

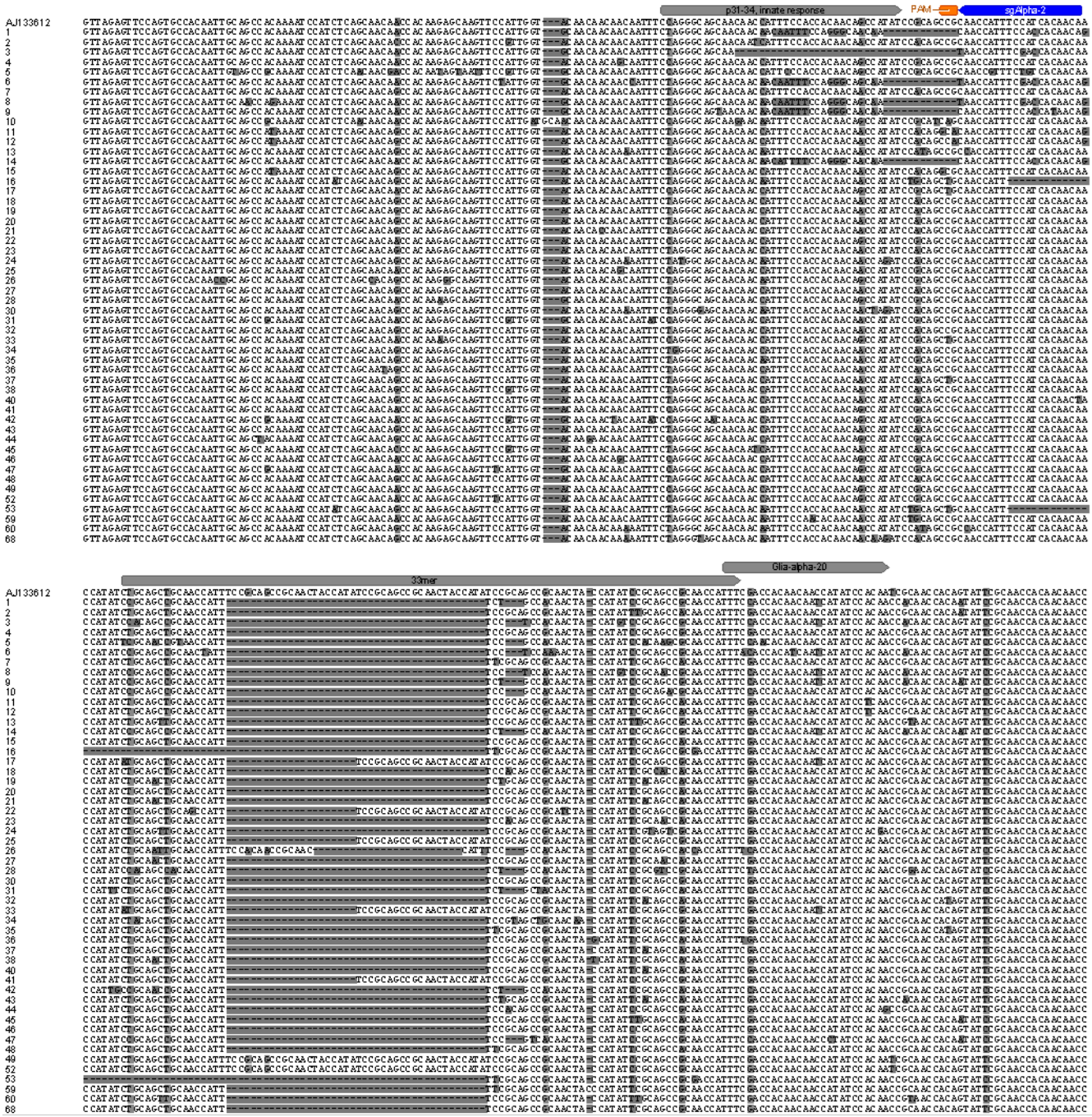
Acknowledgements

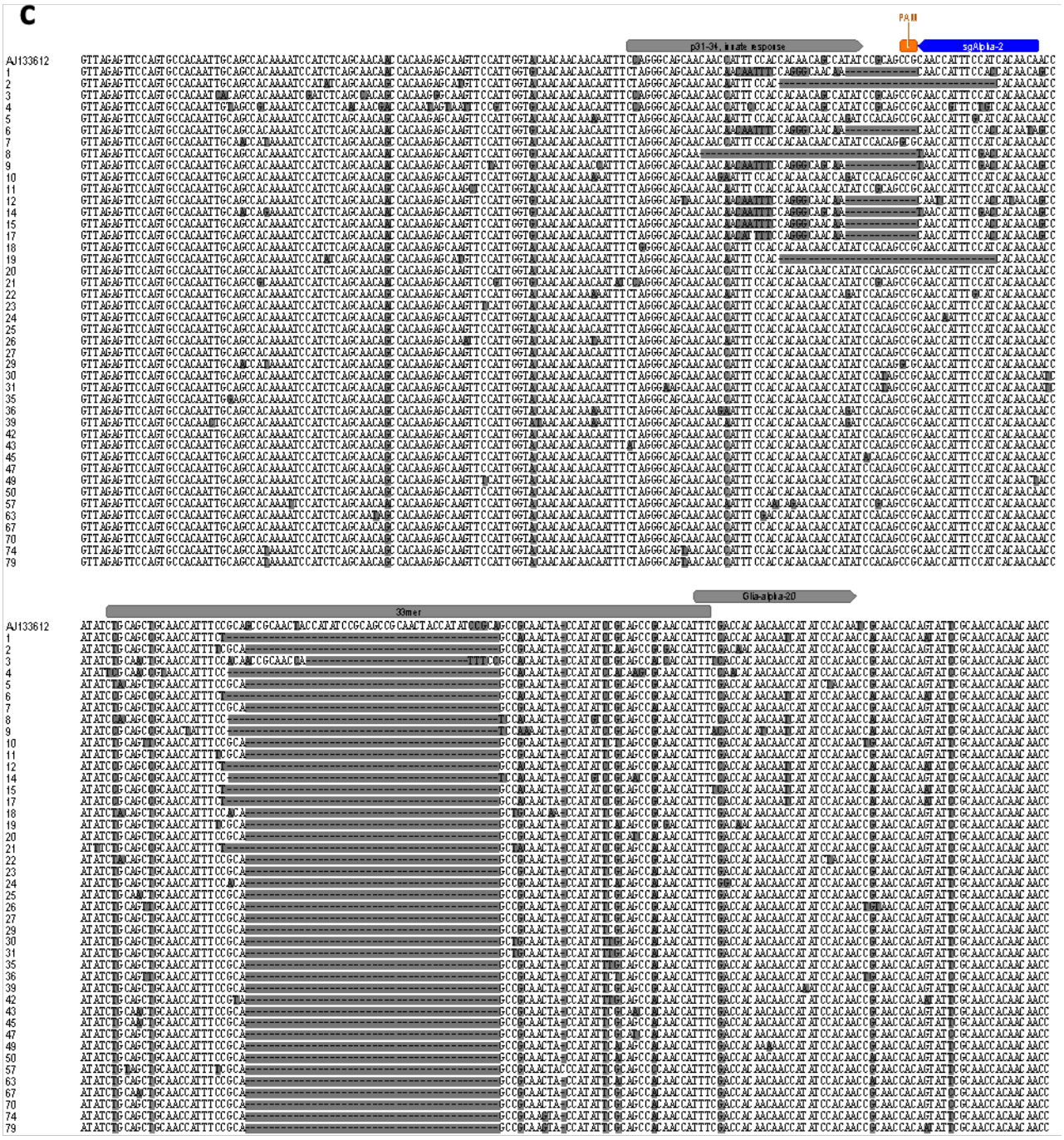
The Spanish Ministry of Economy and Competitiveness (Projects AGL2013-48946-C3-1-R and AGL2016-80566-P) and the European Regional Development Fund (FEDER) supported this work. Javier Gil-Humanes acknowledges the Fundación Alfonso Martín Escudero for its funding support. The technical assistance of Ana García is also acknowledged

Supplementary Material

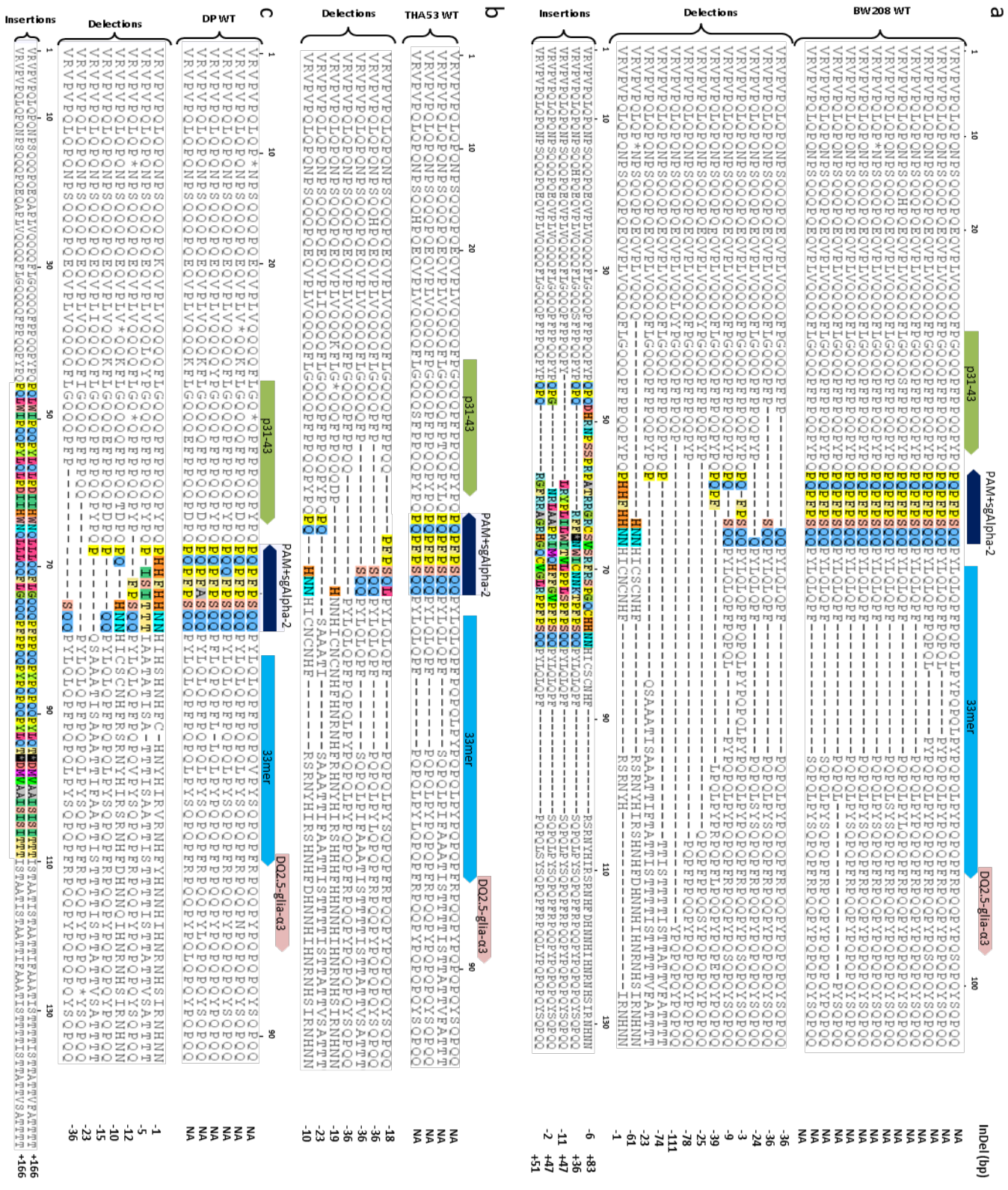


b

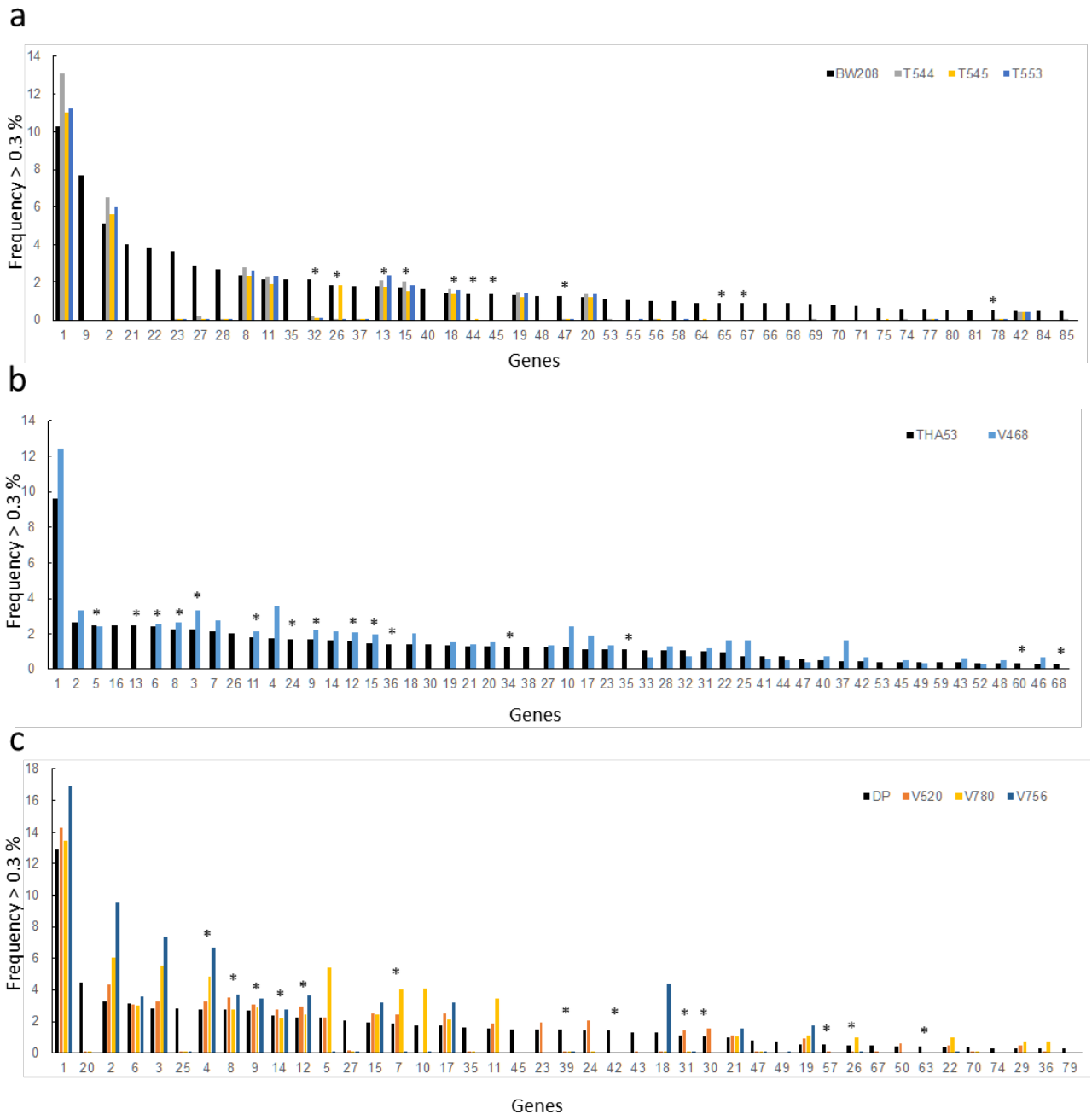

p31-34, innate response
PAM
sgAlpha-2
AJ133612
33mer
Glia-alpha-20



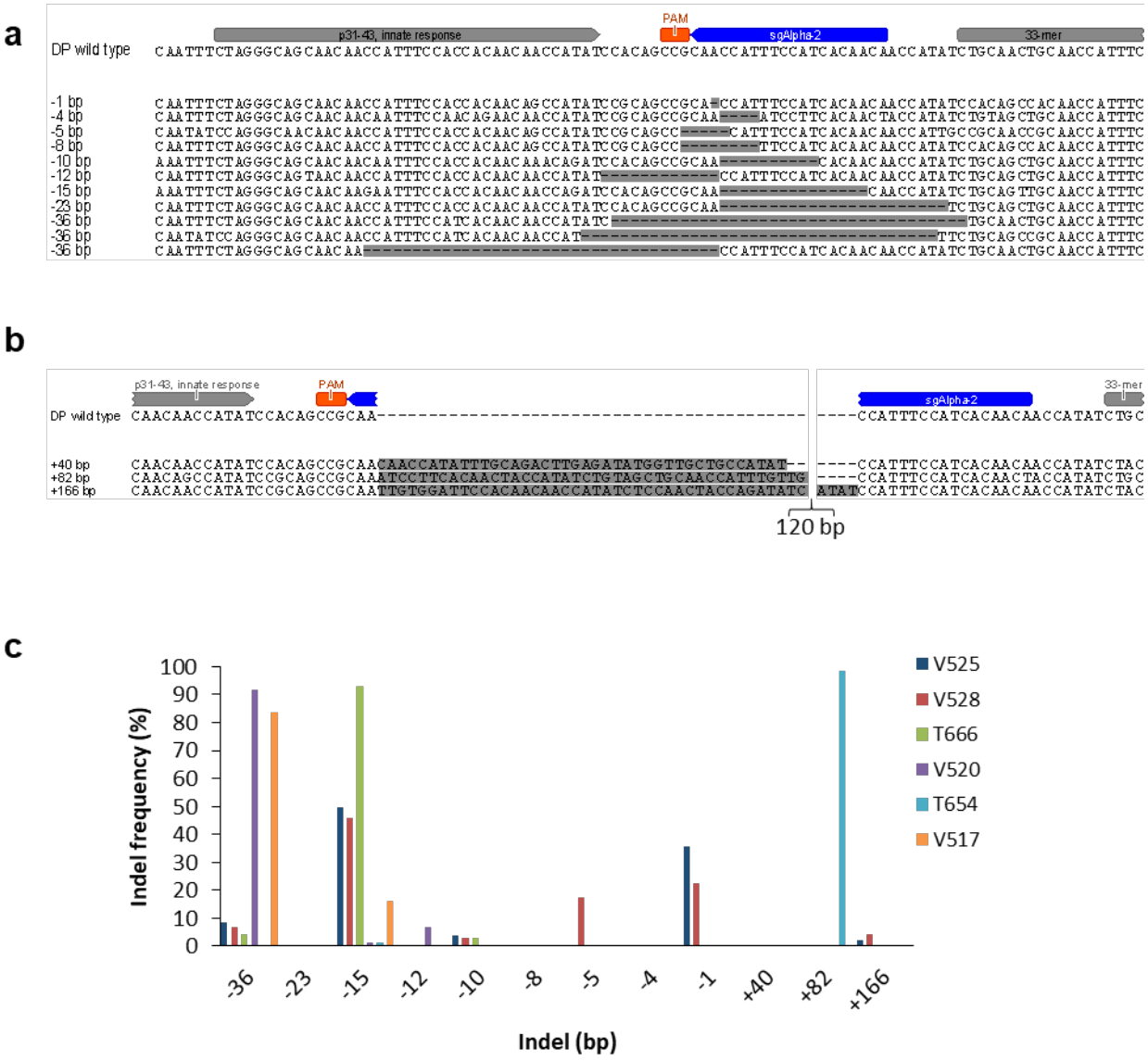
Supplementary figure 1 Nucleotide alignments of the highly represented α-gliadin genes detected by Illumina sequencing (accounting for nearly 85% of the total reads) in the wild type lines of bread wheat (a) cv BW208, (b) cv TAH53, and (c) durum wheat cv DP



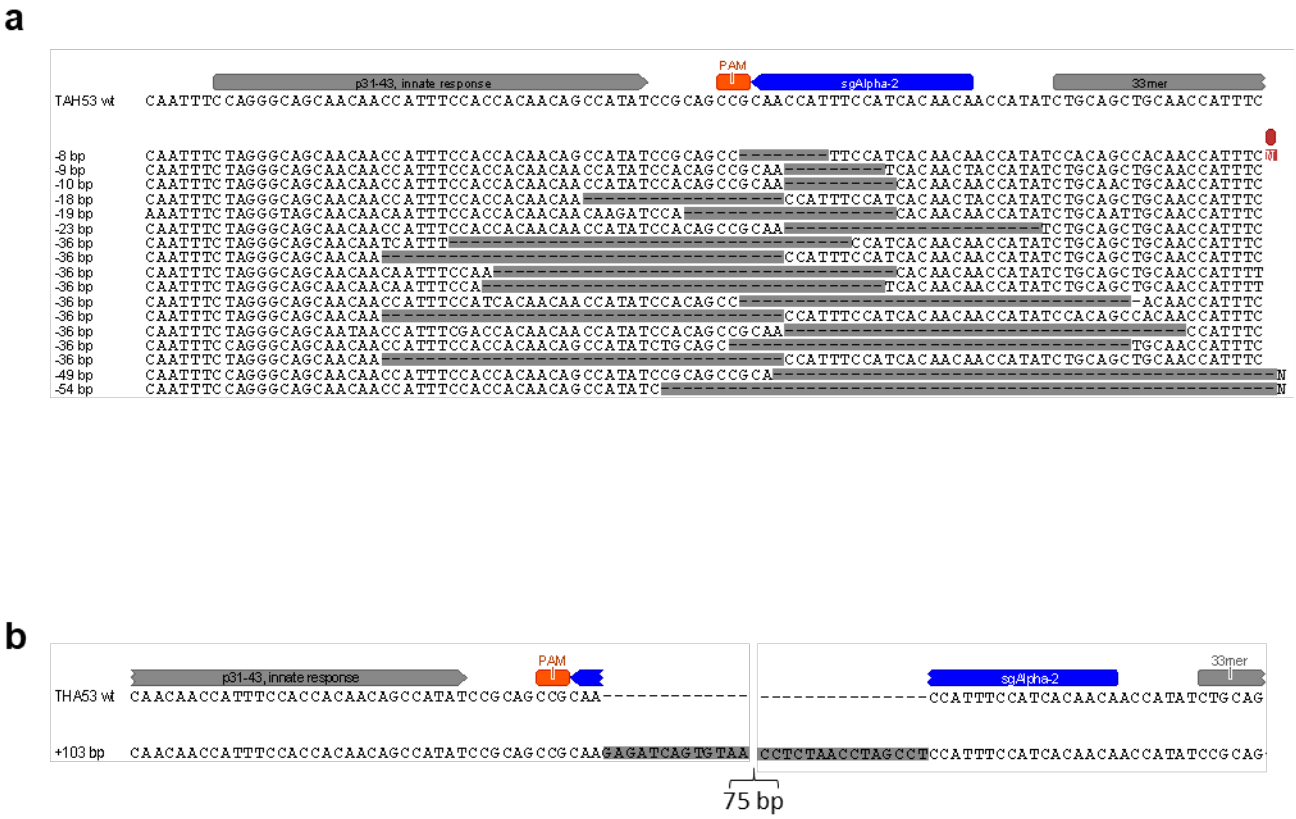
Supplementary figure 2 Protein alignments of the highly represented α-gliadin genes in the wild type lines of bread wheat cv BW208 (a) and cv TAH53 (d), and durum wheat cv DP (f).



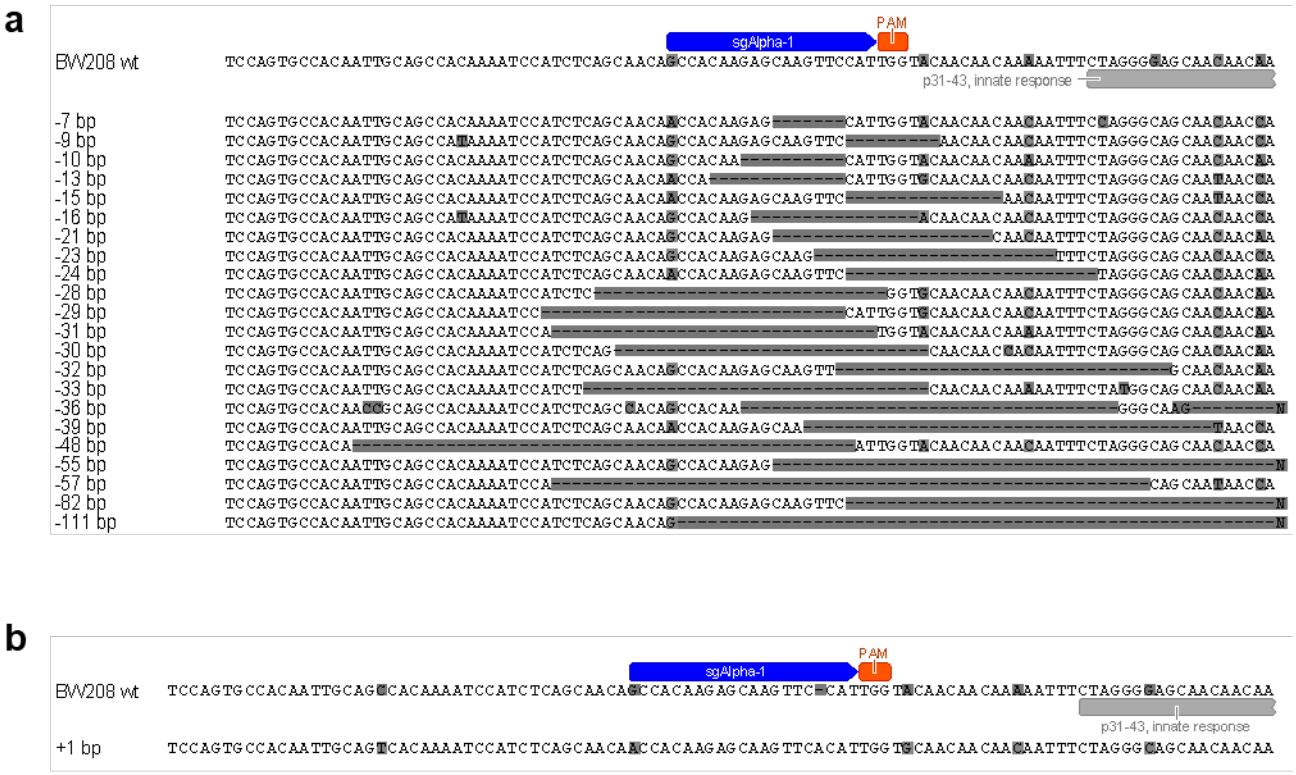
Supplementary figure 3 Estimated α-gliadin genes present in the wild type lines and mutated in the mutant lines by sgAlpha-2.



Supplementary figure 4 Gene editing of α-gliadins in durum wheat cv DP.

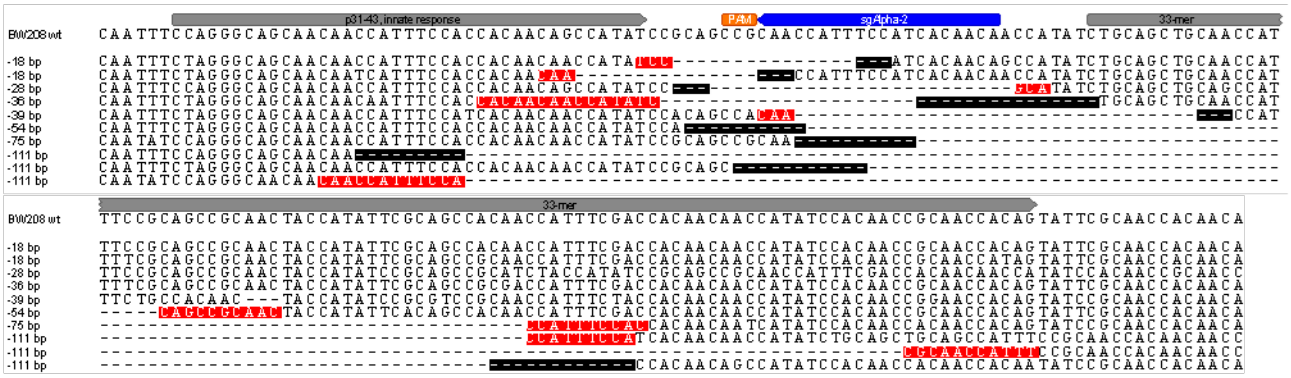


Supplementary figure 5 Gene editing of α-gliadins in bread wheat cv TAH53.

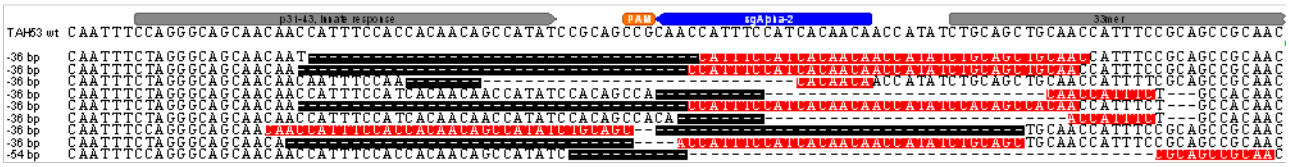


Supplementary figure 6 Gene editing of α-gliadins in bread wheat cv BW208.

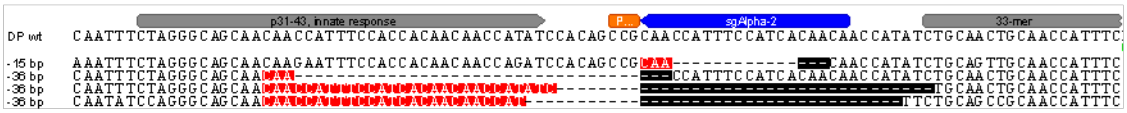
a



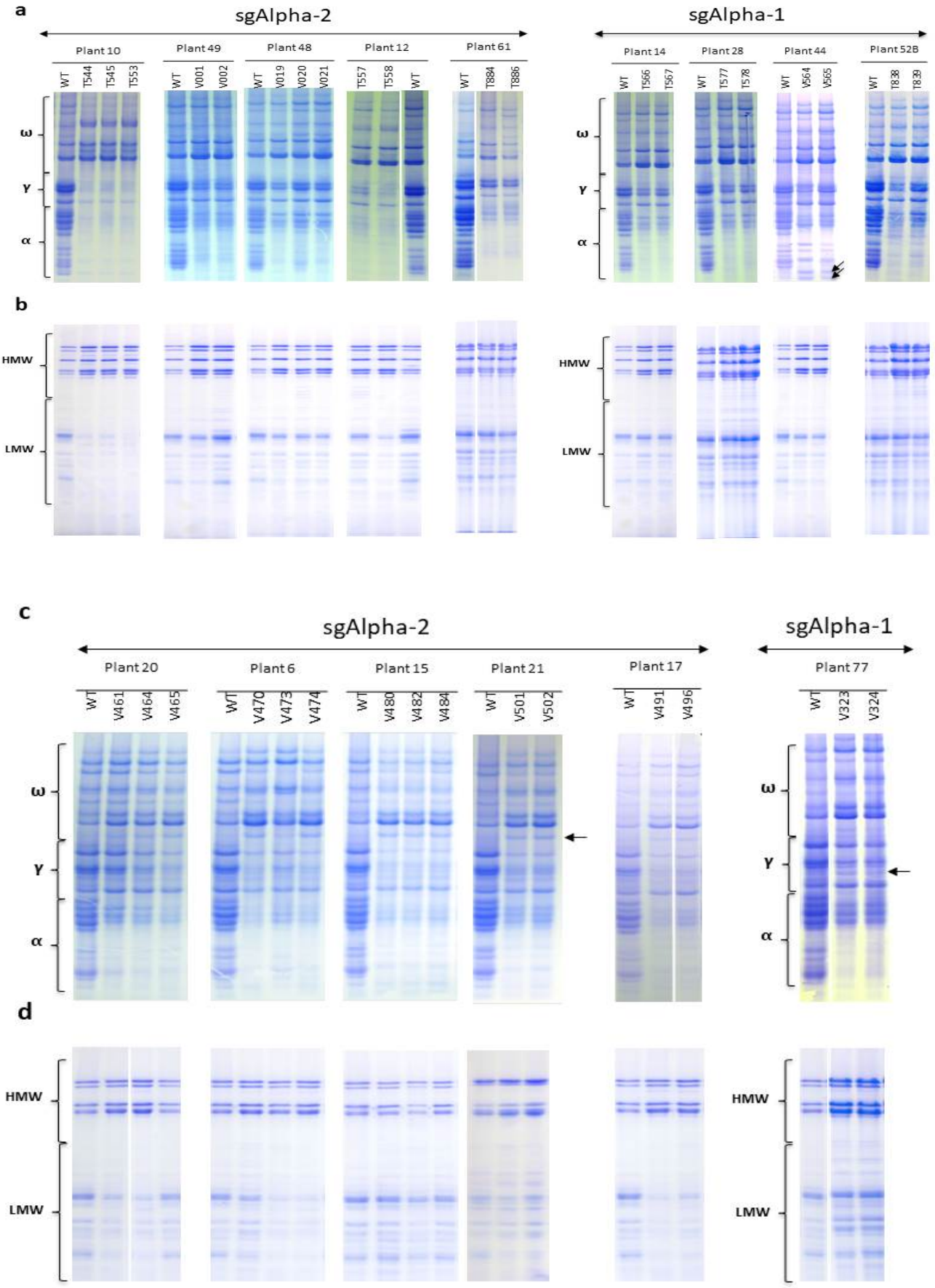
b

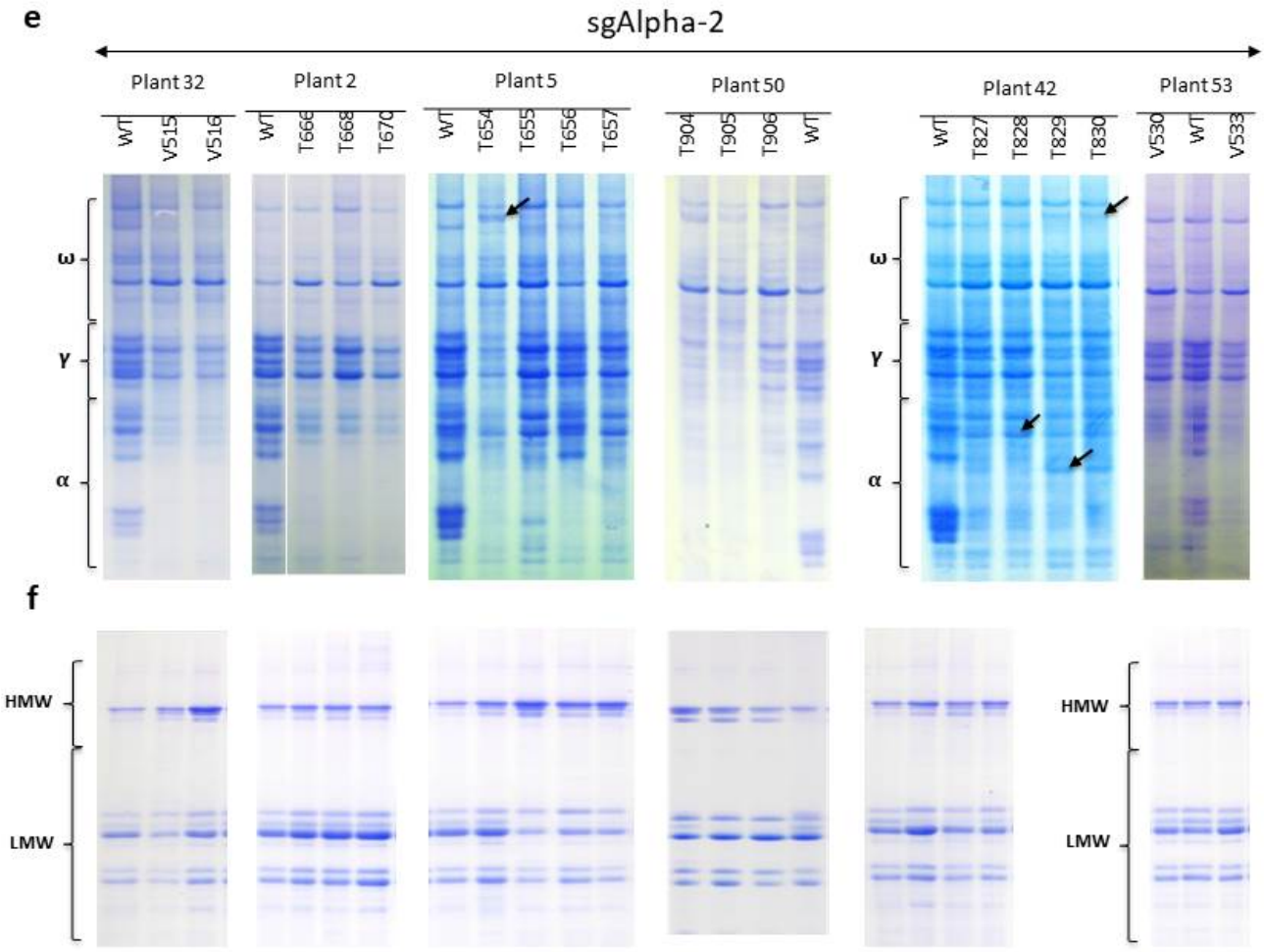


c



Supplementary figure 7 Microhomology-mediated repair of α-gliadins targeted with sgAlpha- 2.


a
sgAlpha-2
sgAlpha-1
Plant 10
Plant 49
Plant 48
Plant 12
Plant 61
Plant 14
Plant 28
Plant 44
Plant 52B
WT
T544
T545
T553
V001
V002
V019
V020
V021
T557
T558
T884
T886
T566
T567
T577
T578
V564
V565
T838
T839
ω
γ
α
b
HMW
LMW
c
sgAlpha-2
sgAlpha-1
Plant 20
Plant 6
Plant 15
Plant 21
Plant 17
Plant 77
V461
V464
V465
V470
V473
V474
V480
V482
V484
V501
V502
V491
V496
V323
V324
d



Supplementary figure 8 Gliadin and glutenin protein fractions analyzed by A-PAGE and SDS-PAGE from T1 half-seeds derived from T0 lines transformed with sgAlpha-1 and sgAlpha-2 constructs.

a

BW208 wt, group 1

ATGAAGACCTTACTCATCCAAACAATCCTCGTGATGGCAATAACCATCGCCACCGCCAATATGCAGGTCGACCCTAGCGGCCAAGTACCATGGCCACAAC
AACAACCATTCCCGCAGCCTCACCAACCATTCTCCCAGCAACCGCAACAAACATTTCCCCAACCCCAACAAACATTCCCCCATCAACCACAACAACAATTTT
CCCAGCCTCAGCAACCACAACAACAATTTATCCAGCCCCAACAACCATTCCCCCAACAACCACAACAAACATATCCCCAGCGACCACAACAACCATTCCCCC
AGACTCAACAACCCCAACAACCATTTCCCCAGTCCCAGCAACCACAACAACCTTTTCCCCAGCCCCAACAACAATTCCCGCAGCCCCAACAACCACAACAAT
CATTCCCCCAACAACAACCATCGTTGATTCAACAATCTCTACAACAACAGTTGAACCCATGCAAGAATTTCCTCTTGCAACAATGCAAACCTGTGTCCTTGG
TGTCATCCCTCTGGTCAATGATCTTGCCACGAAGCGATTGCCAGGTGATGCGGCAACAATGTTGCCAACAACTAGCACAAATTCCTCAGCAACTCCAGTGT
GCAGCCATCCATAGCATCGTGCATTCCATCATCATGCAGCAAGAACAACAAGAACAACGACAGGGTGTGCAAATCCTGGTGCCACTGTCTCAACAGCAAC
AGGTAGGTCAAGGTACTCTCGTCCAAGGTCAGGGCATCATCCAACCTCAACAACCAGCTCAATTGGAGGTGATTAGGTCATTGGTGTTGCAAACTCTTGC
AACCATGTGCAACGTGTATGT

BW208 wt, group 2

ATGAAGACCTTACTCATCCTAACAATCCTTGCGATGGCAACAACAATCGCCACTGCCAATATGCAGGTCGACCCTAGCAGCCGAGTACAATGGCCACAAG
AACAACCACCCCCCCAGTCCCAACAACCATTCTCCCAGCAACCACAACAAATATTTCCCCAACCCCAACAAACATTCCCCCATCAACCACAACAAGCATTTC
TCCAACCTCAACAAACATTCCCCCGTCGACCACAACAACAATTTCCCCAGCCCCAGCAACCACAACAACCATTTCCCTCAGCCCCAACAACCCCAACTACCAT
TTCCCCAACAACCACAACAACCATTCCCCCAGCCTCAACAACCCCAACAACCATTTCCCCAGTCACAGCAACCACAACAACCTTTTCCCCAGCCCCAACAAC
AATTTCCGCAGCCCCAACAACCACAACAATCATTCCCCCAACAACAACAATGGATGATTCAGTCATTTCTACAACAACAGATGAACCCCTGCAAGAATTTC
CTCTTGCAGCAATGCAACCCTGTGTCATTGGTGTCATCTCTCGTGTCAATAATCTTGCCACGAAGTGATTGCCAGCTGATGCAGCAACAATGTTGCCAACA
ACTAGCACAAATTCCTCAACAACTCCAGTGCGCAGCCATCCACAACGTCGCGCATTCCATCATCATGCAGCAAGAACAACAACGAGGCGTGCAGATCCTG
CGGCCACTATTTCAGCTCGCCCAGGGTCTGGGTATCATCCAACCTCAACAACCAGCTCAATTGGAGGGGATCAGGTCATTGGTATTGAAAACTCTTCCAAC
CATGTGCAACGTGTATGT

BW208 wt, group 3

ATGAAGACCTTACTCATCCTGACAATCATTGCGGTGGCACTAACTACCACCACCGCCAATATACAGGTCGACCCTAGTGGCCAAGTACAATGGCCACAAC
AACAACAACCATTCCCCCAGCCCCAACAACCACAACAAATTTTTCCCCAACCCCAACAAACATTCCCCCATCAACCACAACAAGCATTTCCCCAACCCCAAC
AAACATTCCCCCATCAACCACAACAACAATTTCCCCAGCCCCAGCAACCACAACAACCATTTCCCCAGCCCCAACAAC
CACAACAACCATTTCCCCAGCAACCACAACAACAATTTCCCCAGCCCCAACAACCACAACAACCATTTCCCCAGCCCCAACAACCCCAACGTACCATTTCCGGC
AACAACCACAACAACCATTCCCCCAGCCTCAACAACCCCAACAACCATTTCCCCAGTTACAGCAACCACAACAACCTTTACCCCAGCCCCAACAACCGCAAC
AACCATTCCCCCAGCAACAACAACCATTGATTCAGCCATACCTACAACAACAGATGAACCCCTGCAAGAATTACCTCTTGCAACAATGCAACCCTGTGTCA
TTGGTGTCATCCCTCGTGTCAATGATCTTGCCACGAAGTGATTGCAAGGTGATGCGGCAACAATGTTGCCAACAACTAGCACAGATTCCTCAGCAGCTCCA
GTGCGCAGCCATCCATGGCGTCGTGCATTCCATCATCATGCAGCAAGAACGACAACAACAACAACAACAACAACAAGGCATACAGATCATGCGGCCACT
ATTTCAGCTCGTCCAGGGTCAGGGCATCATCCAACCTCAACAACCAGCTCAATTGGAGGTGATCAGGTCATTGGTATTGGGAACTCTTCCAACCATGTGCA
ACGTGTATGT

BW208 wt, group 4

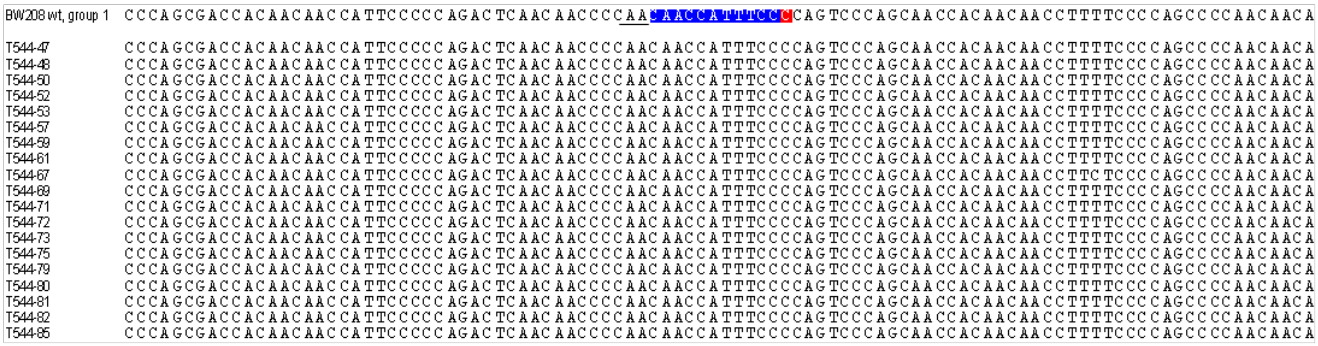
ATGAAGACCTTACTCATCCWAACAATCCTYGYGATGGCAATAACCATCGGCACCGCCAATATSCAGGTCGACCCTAGCRGCCAAGTACAATGGCYACAAC
AACAACYAGTCCCMCAGCYYCAMCARCCATTMTCCCAGCAACCACAACAAACATTTCCCCRACCYCAACAAACATTCCCCCATCAACCACAACAACAAKTT
YCCCAGCCTCAGCAACCACAACAACMATTTCTCCAGCCCCGACAACCATTCCCCCAACAACCACAACAACCATATCCCCAGCAACCACAGCAACCGTTCCC
CCAGACTCAACAACCCCAACAACCATTTCCCCAGTCCAAGCAACCACAACAACCTTTTCCCCAGCCCCAACAACCGCAACAATCATTCCCCCAACAACAACC
ATCGTTGATTCAACAATCTCTACAACAACAGTTGAACCCATGCAAGAATTTCCTCTTGCAGCAATGCAAACCTGTGTCCTTGGTGTCATCSCTCTGGTCAAT
CATCTTGCCACCAAGCGATTGCCAGGTGATGCGGCAACAATGTTGTCAACAACTAGCACAAATTCCTCAGCAACTCCAGTGTGCAGCCATCCATAGCGTC
GTGCATTCCATCATCATGCAGCAAGAACAACAAGAACAACTACAGGGTGTGCAAATCCTGGTGCCACTGTCTCAACAGCAACAAGTGGGTCAAGGTATTC
TCGTCCAGGGTCAAGGCATCATCCAACCTCAACAACCAGCTCAATTGGAGGTGATCAGGTCATTGGTGTTGCAAACTCTTCCAACCATGTGCAACGTGTAT
GT

BW208 wt, group 5

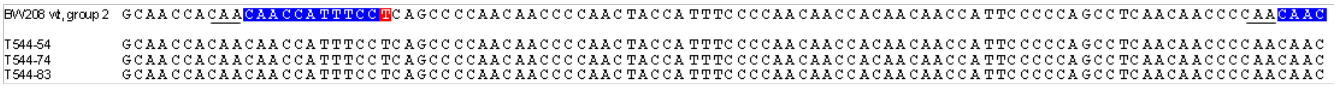
ATGAAGACCTTACTCATCCTGACAATCCTTGCGATGGCAATAACCATCGGCACCGCCAATATCCAGGTCGACCCTAGCGGCCAAGTACAATGGCTACAAC
AACAACTAGTCCCCCAGCTCCAACAGCCATTATCCCAGCAACCACAACAAACATTTCCCCAACCTCAACAAACATTCCCCCATCAACCACAACAACAAGTTC
CCCAGCCTCAGCAACCACAACAACCATTTCTCCAGCCCCAACAACCATTCCCCCAACAACCACAACAACCATTCCCCCAGACTCAACAACCACAACAACCAT
TTCCCCAGCAACCACAACAACCATTTCCCCAGACTCAACAACCCCAACAACCATTCCCCCAACAACCACAACAACCATTCCCCCAGACTCAACAACCCCAACC
AACCATTTCCCCAGCTCCAGCAACCACAACAACCTTTTCCCCAGCCCCAACAACAATTGCCGCAGCCCCAACAACCGCAACAATCATTCCCCCAACAACAAC
GGCCATTCATTCAACCATCTCTACAACAACAGTTGAACCCATGCAAGAATATCCTCTTGCAACAATGCAAACCTGCGTCATTGGTGTCATCCCTCTGGTCAA
TAATCTGGCCACAAAGCGATTGCCAAGTGATGCGGCAACAATGCTGCCAACAACTAGCACAGATTCCTCAACAGCTCCAGTGCGCAGCCATCCATAGCGT
CGTGCATTCCATCATCATGCAGCAGCAGCAGCAACAACAACAACAACAAGGCATGCATATCTTTCTGCCACTATCTCAGCAGCAACAGGTGGGTCAAGGT
TCTCTAGTCCAAGGCCAGGGCATCATCCAACCACAACAACCAGCTCAATTGGAGGCGATCAGATCATTGGTGTTGCAAACTCTTCCATCCATGTGCAACGT
GTATGT

b

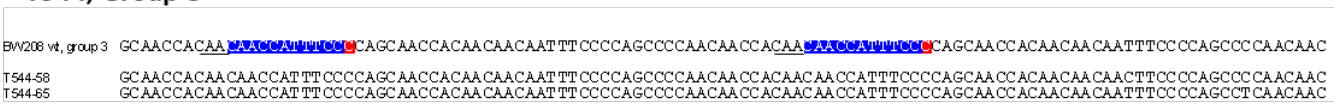
T544, Group 1

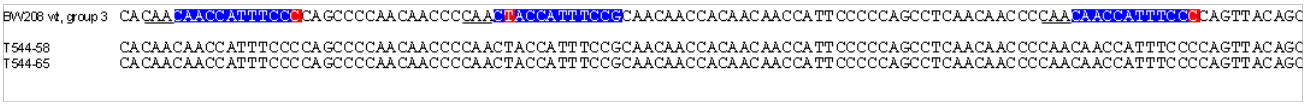


T544, Group 2

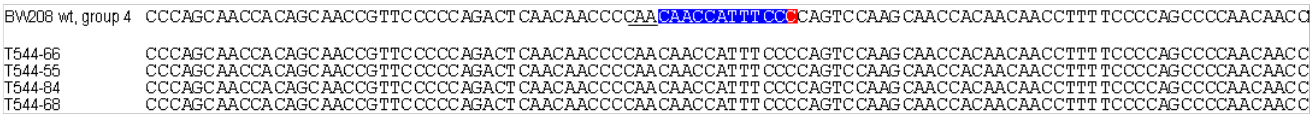


T544, Group 3



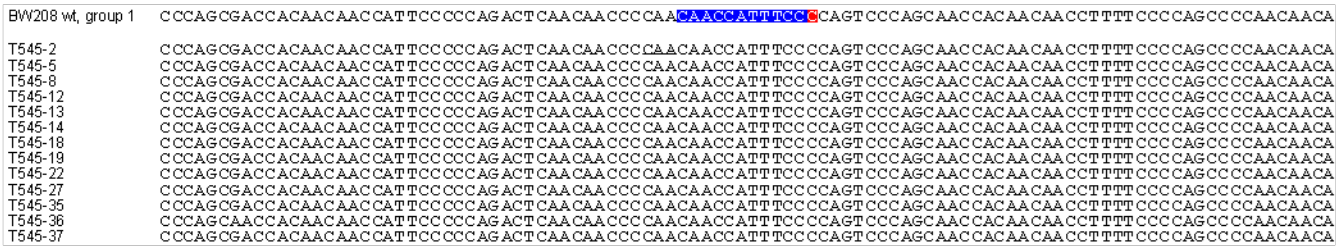


T544, Group 4

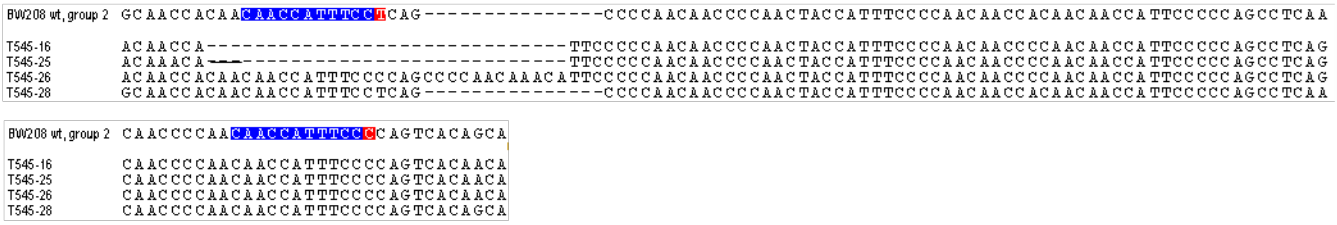


C

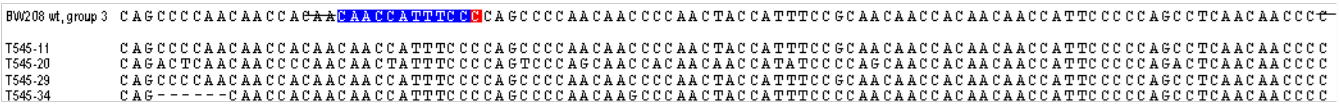
T545, Group 1

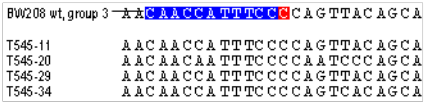


T545, Group 2

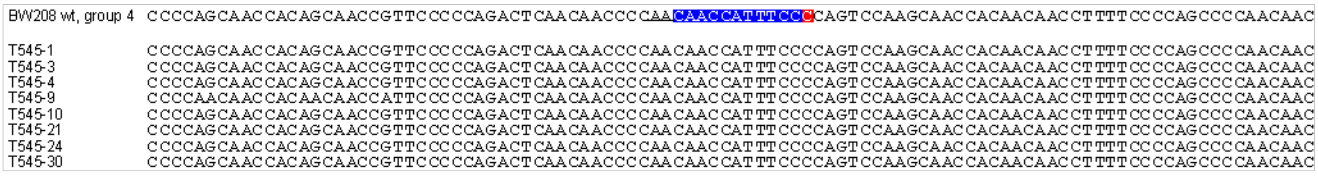


T545, Group 3

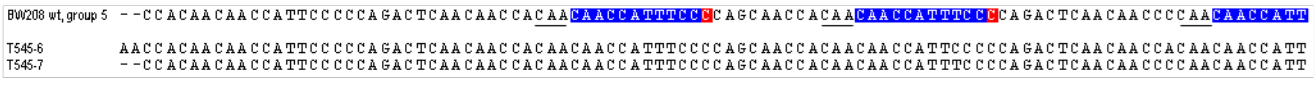


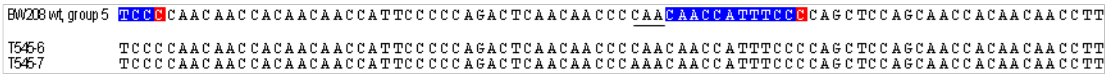


T545, Group 4



T545, Group 5





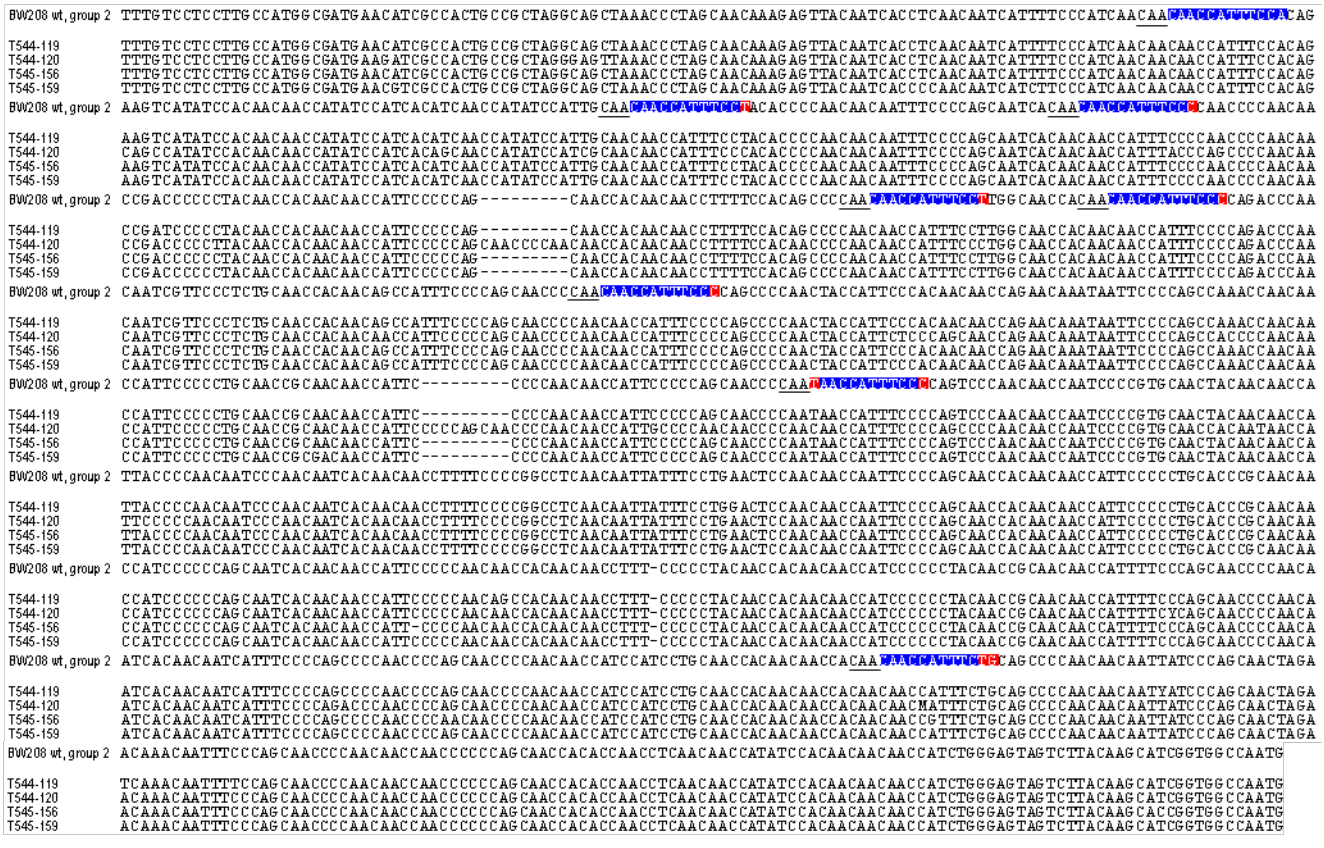
Supplementary figure 9 Off-target mutations detection in γ-gliadin genes of BW208 wild type and two T1 mutant lines.

a

Group 1

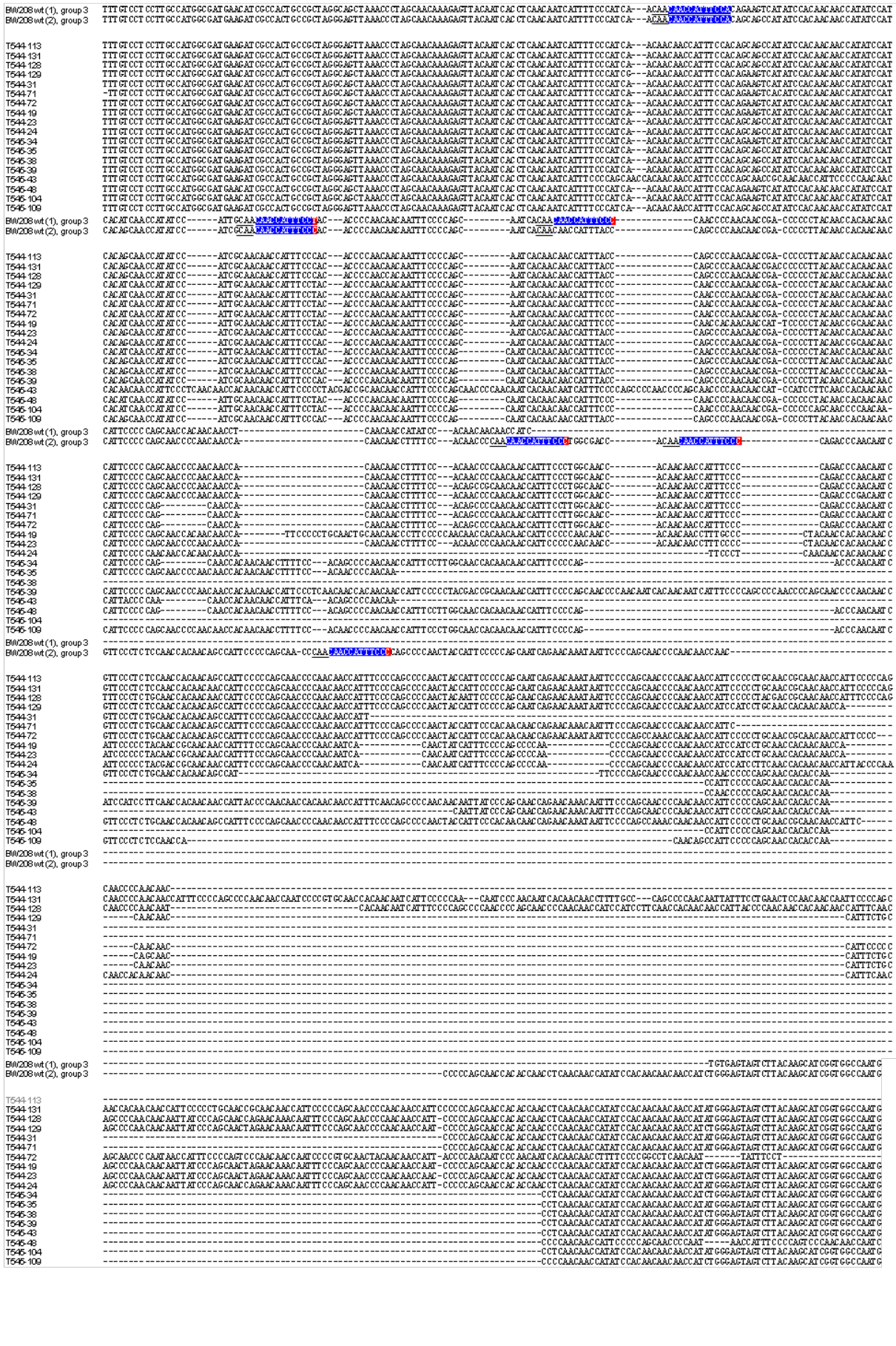
b

Group 2

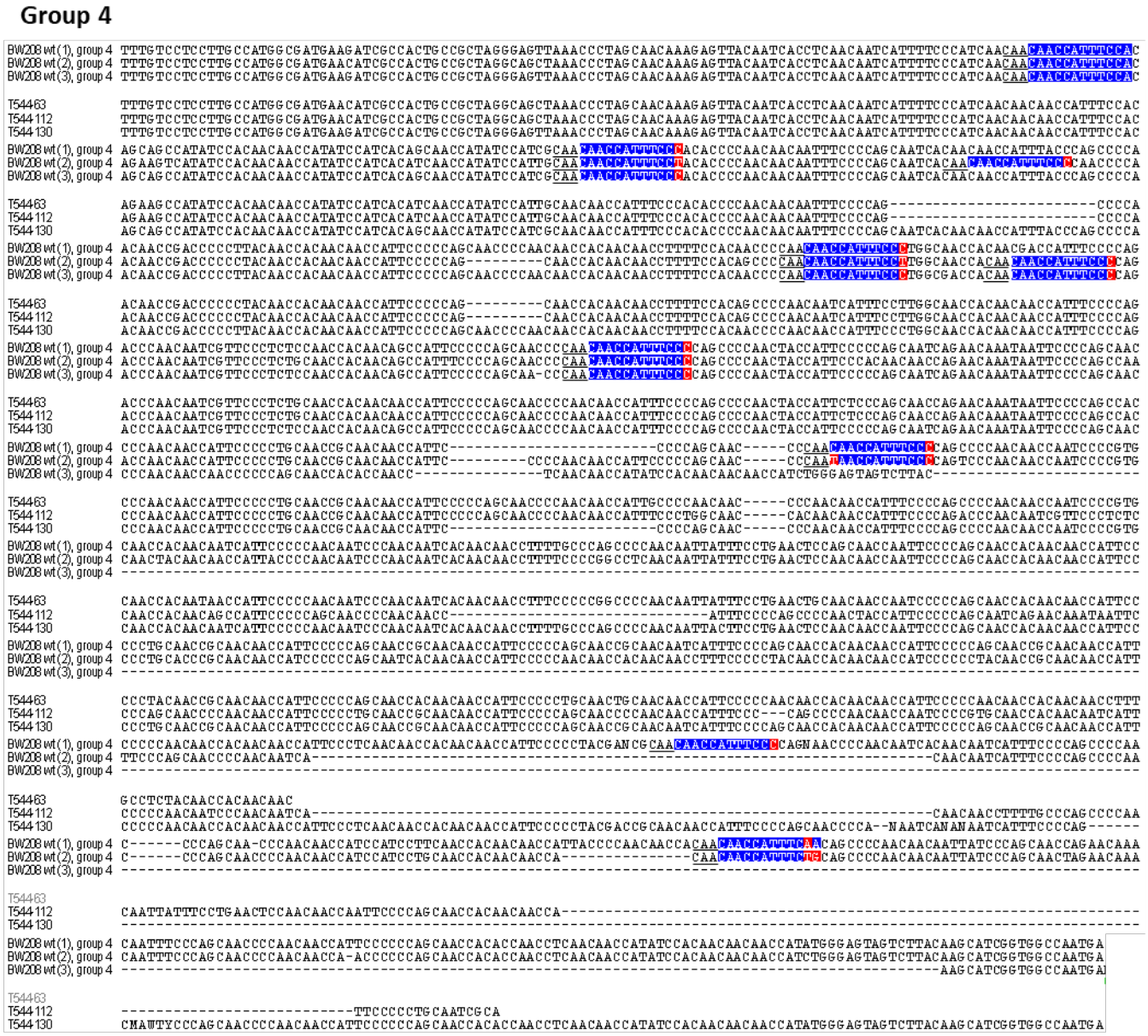


C

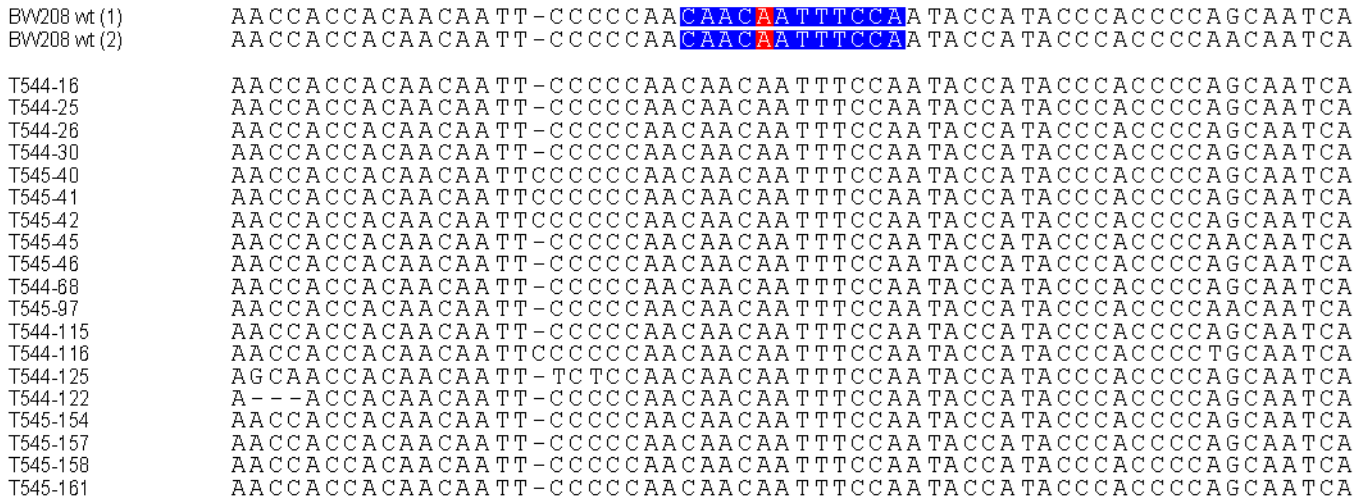
Group 3


BW208 wt (1), group 3
BW208 wt (2), group 3
T544-113
T544-131
T544-128
T544-129
T544-31
T544-71
T544-72
T544-19
T544-23
T544-24
T545-34
T545-35
T545-38
T545-39
T545-43
T545-48
T545-104
T545-109

d



Supplementary figure 10 Off-target mutations detection in ω1,2-gliadin genes of BW208 wild type and two T1 mutant lines.



Supplementary figure 11 Off-target mutations detection in ω5-gliadin genes of BW208 wild type and two T1 mutant lines (T544 and T545).

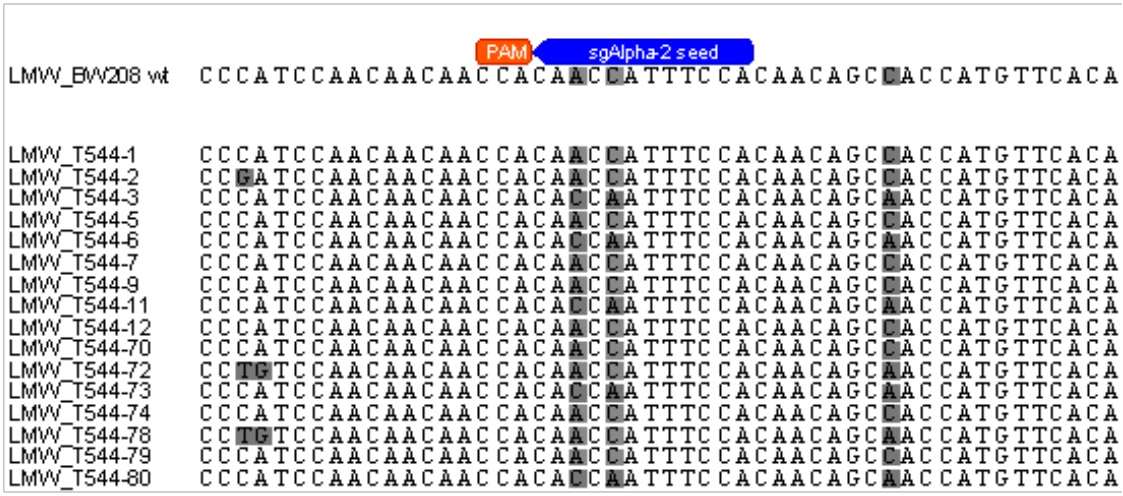
a

Prolamin off-target

	No. of analyzed sequences	No. of off-target hits minimal sequence sgRNA+PAM*	
		sgAlpha-1	sgAlpha-2
α-gliadins	156	152	153
γ-gliadins	179	0	0
ω-gliadins	15	0	0
HMW-glutenins	40	0	0
LMW-glutenins	239	0	71

***up to one mismatch in the seed sequence + perfect PAM**

b

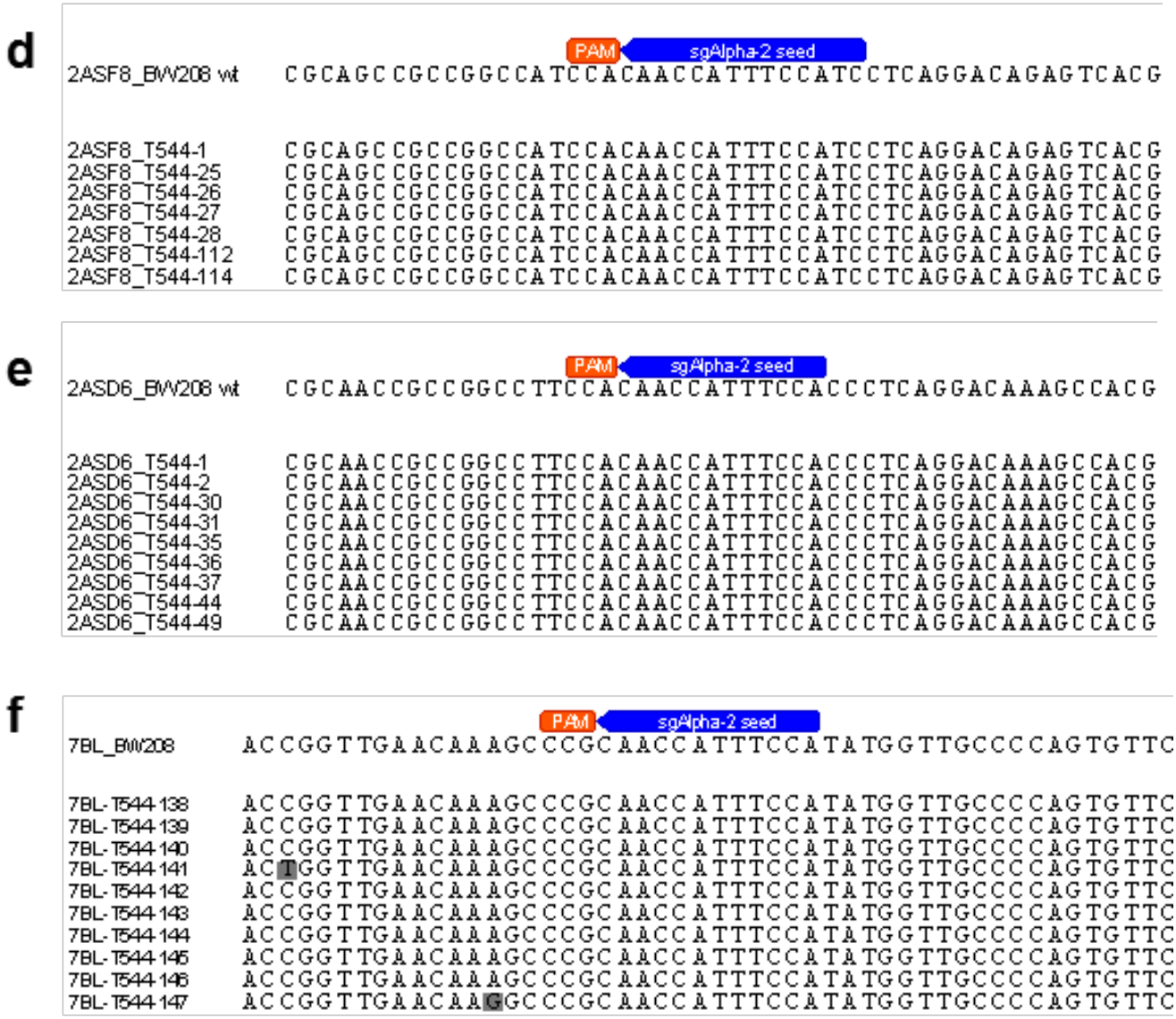


C

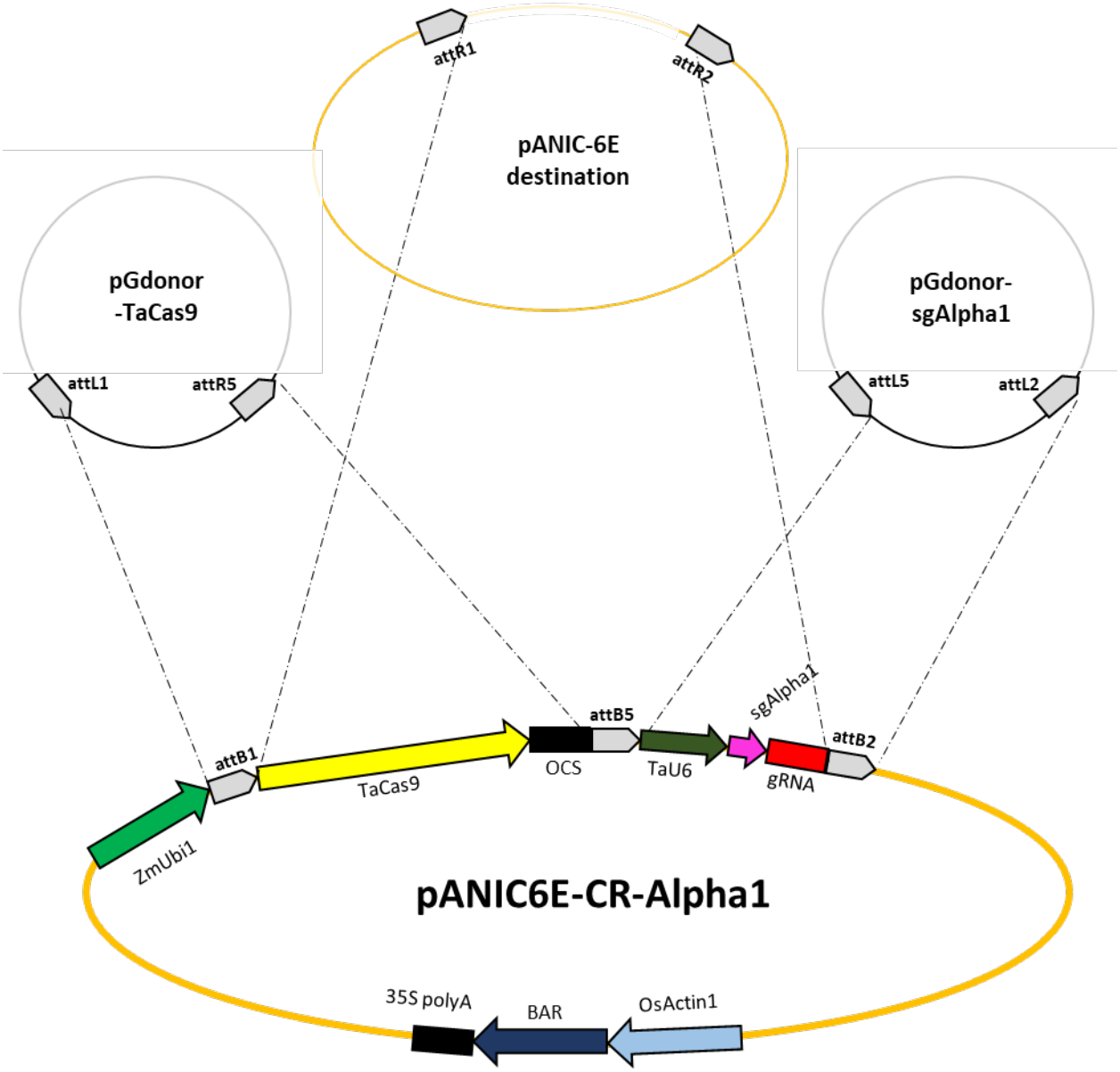
Non-prolamin off-target

sgRNA*	Off-target hits	Annotated genes	Description	No. of clones sequenced	No. of clones with mutations
sgAlpha-1	41	Traes_4AL_4FF5B8837	Alpha-gliadin pseudo gene		
sgAlpha-2	50	Traes_4AL_4FF5B8837	Alpha-gliadin pseudo gene	NA	-
		Traes_2AS_8FCC59363	Uncharacterized	24	0
		Traes_2AS_D659E88E9	Uncharacterized	28	0
		Traes_7BL_F621D9B9E	MADS box transcription factor	30	0

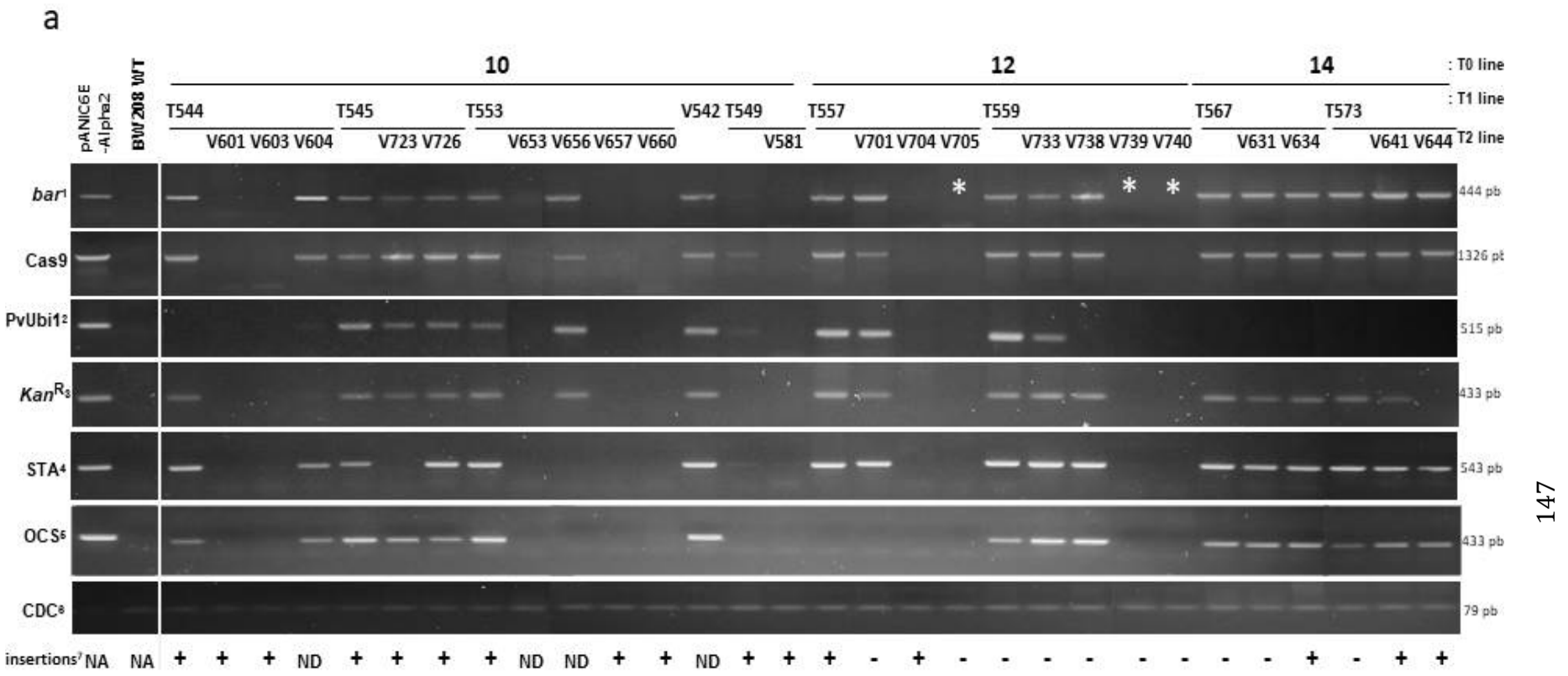
***** Perfect match seed sequence (12 nt) + PAM**



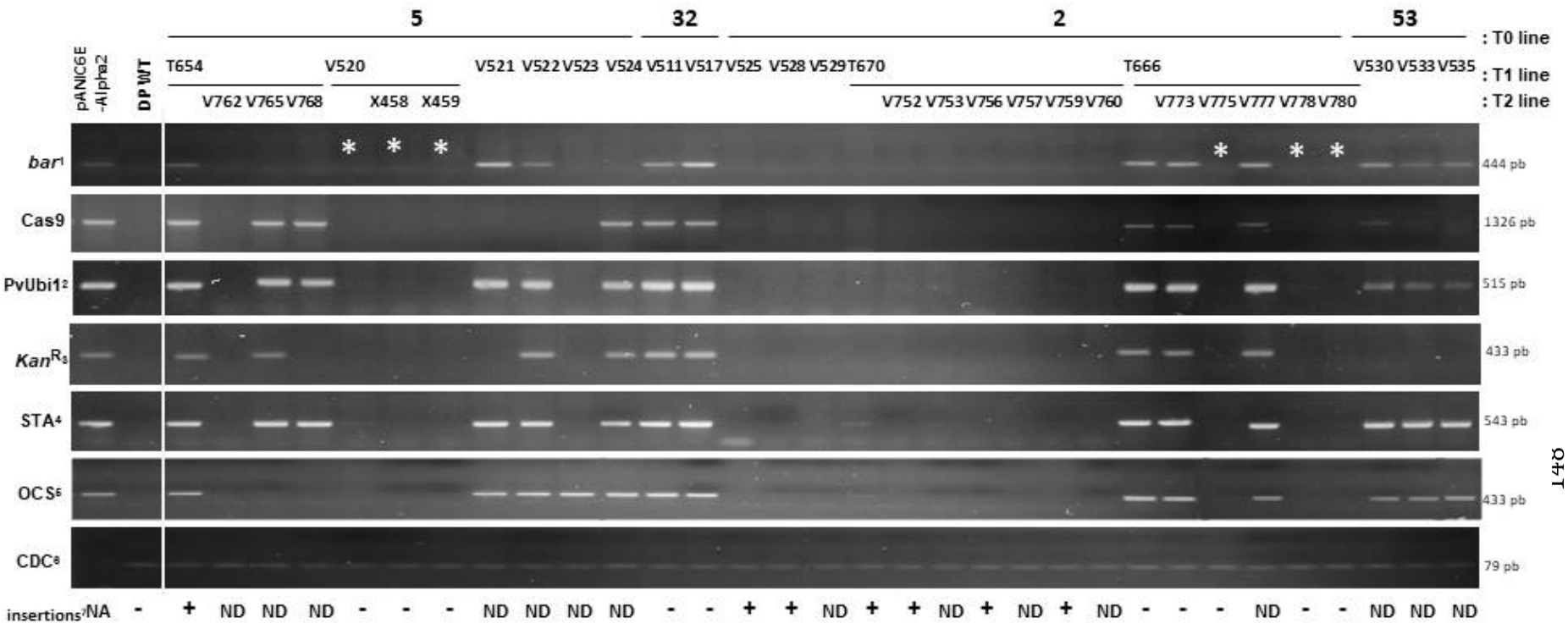
Supplementary figure 12 Off-target mutations detection in BW208 mutant lines.

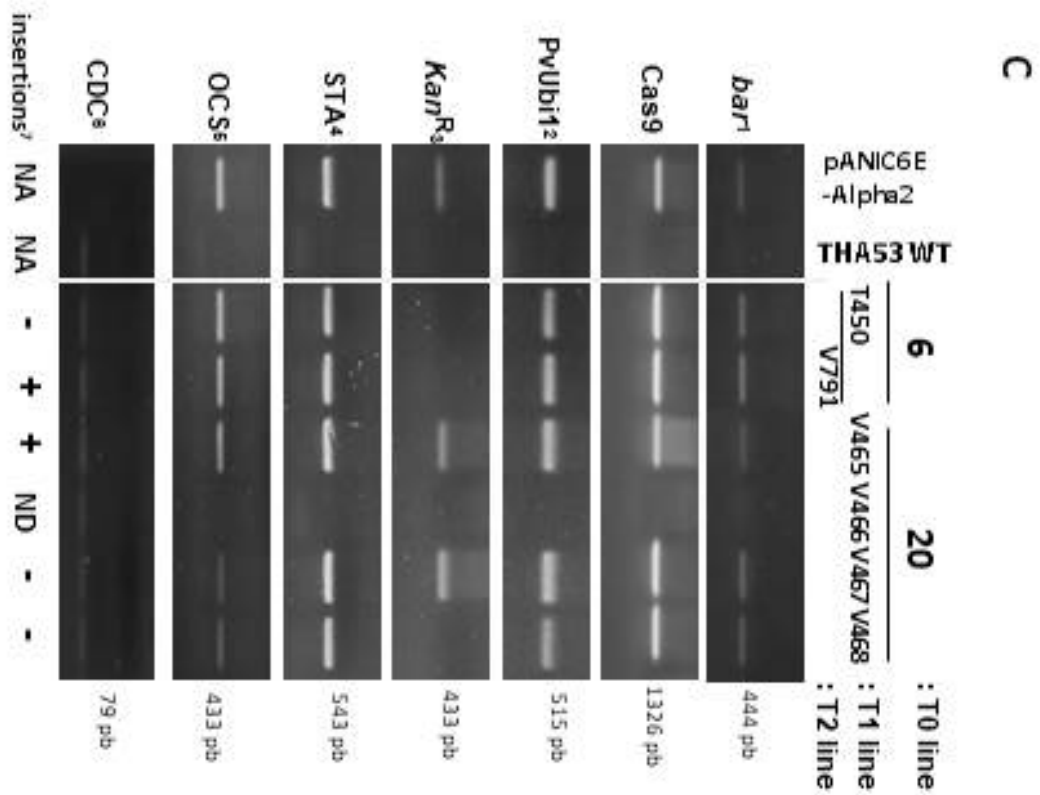


Supplementary figure 13 Multisite Gateway cloning of pANIC6E-CR-Alpha1 vector.

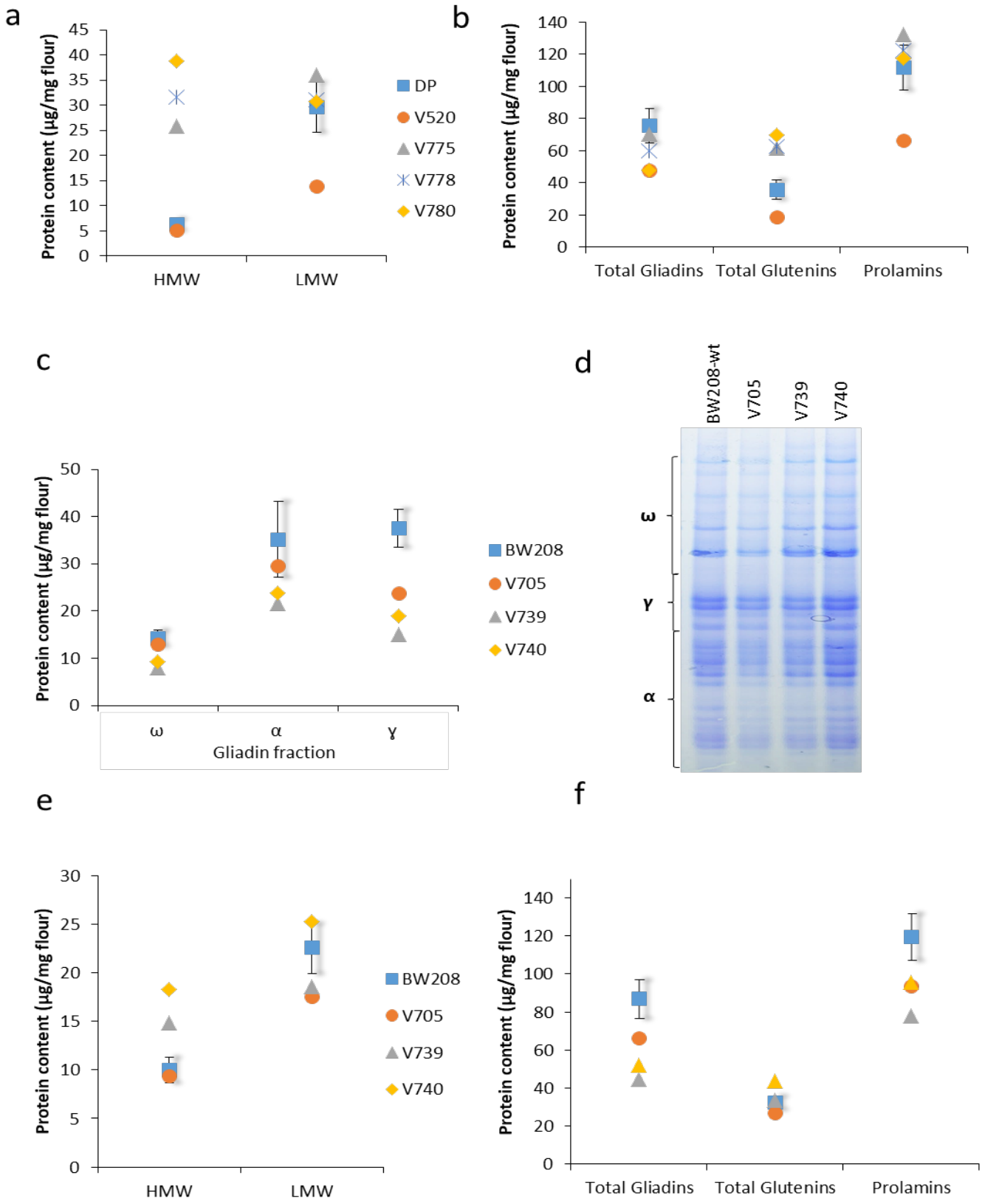

a
pANIC6E -Alpha2
BW208 WT
10
12
14
: T0 line
: T1 line
T544
T545
T553
V542 T549
T557
T559
T567
T573
V601 V603 V604
V723 V726
V653 V656 V657 V660
V581
V701 V704 V705
V733 V738 V739 V740
V631 V634
V641 V644
T2 line
bar1
Cas9
PvUbi12
KanRs4
STA4
OCS6
CDC8
444 pb
1326 pb
515 pb
433 pb
543 pb
433 pb
79 pb
insertions7 NA NA + + + ND + + + + ND ND + + ND + + + - + - - - - - - - + - + +

b


5
32
2
53
: T0 line
: T1 line
: T2 line
pANIC6E -Alpha2
DP WT
T654
V520
V521 V522 V523 V524 V511 V517 V525 V528 V529
T670
T666
V530 V533 V535
V762 V765 V768
X458 X459
V752 V753 V756 V757 V759 V760
V773 V775 V777 V778 V780
bar1
Cas9
PvUbi12
KanR5
STA4
OCS6
CDC8
444 pb
1326 pb
515 pb
433 pb
543 pb
433 pb
79 pb
insertions7NA - + ND ND ND - - - ND ND ND ND - - + + ND + + ND + ND + ND - - - ND - - ND ND ND



Supplementary figure 14 Analysis by PCR and Illumina high-throughput sequencing for the presence of the plasmid DNA; *bar* and Cas9 genes, PVS1 stability (sta) region, Octopine synthase polyA signal, and *Panicum virgatum* ubiquitin 1 promoter; and insertions in sgAlpha-2 derived lines.



Supplementary figure 15 Protein analysis of non-transgenic transgenic (transgene-free and insertion-free) lines determined by RP-HPLC and A-PAGE gels.

Supplementary table 1. List and sequence of primers for PCR and Illumina sequencing

name	sequence (5'-3')	Description
aGli900 F1	GTTAGAGTTCCAGTGCCACAA	Forward primer for amplicon deep sequencing
33mer1R2_Ok	GGTTGTTGTGGTTGCGRATA	Reverse primer for amplicon deep sequencing
BAR1F	GTCTGCACCATCGTCAACC	Forward primer for detection of *bar* gene
BAR2R	GAAGTCCAGCTGCCAGAAAC	Reverse primer for detection of *bar* gene
fJG218F	TAAGGTCCTCTCCGCCTACA	Forward primer for detection of Cas9
fJG218R	GGCGGTAAGGATCTGAGCTA	Reverse primer for detection of Cas9
M13 Reverse	TCACACAGGAAACAGCTATGAC	Forward primer for sequencing with pGEM®-T Easy Vector system
M13 Forward	CGCCAGGGTTTTCCCAGTCACGAC	Reverse primer for sequencing with pGEM®-T Easy Vector system
KanR_Fw	TTATGCCTCTTCCGACCATC	Forward primer for detection of kanamycin resistance gene
KanR_Rv	ATTCCGACTCGTCCAACATC	Reverse primer for detection of kanamycin resistance gene
PvUbi1_Fw	CGTCTCGCAAAATAGCACAA	Forward primer for detection of *Panicum virgatum L*. ubiquitin promoter
PvUbi1_Rv	CATAATCGATCGAGGGGAGA	Reverse primer for detection of *Panicum virgatum L*. ubiquitin promoter
STA_Fw	GCTGCGTATATGATGCGATG	Forward primer for detection of stability region from PVS1 plasmid
STA_Rv	GACTCAAGAATGGGCAGCTC	Reverse primer for detection of stability region from PVS1 plasmid
OCS_Fw	CCGGTTTCGGTTCATTCTAA	Forward primer for detection of octopine synthase polyA signal
OCS_Rv	GTTGAATGGTGCCCGTAACT	Reverse primer for detection of octopine synthase polyA signal
Primers for detection of insertions in T1 and T2 plants (all with forward primer aGli900F1)		
pJG001	ATCGATCGTGGTGTCGAAGT	Reverse primer for detection of +158 bp insertion
pJG002	TGATGGAAATGGCGTTTTATTATTACA	Reverse primer for detection of +36 bp insertion
pJG003	ACTCAAAGGCGGTAATACGGT	Reverse primer for detection of +47 bp insertion
pJG004	TCGTAATCCCACACACTGGC	Reverse primer for detection of +51 bp insertion
pJG005	ACTAACAGAACATCGGCCCC	Reverse primer for detection of +83 bp insertion
Primers for cloning and sequencing of possible off-target		
pJG006	AGACCTTCCTCATCTTTGC	Forward primer for cloning of LMW for sequencing
pJG007	TGCTGCGATAATGGTGGTTG	Reverse primer for cloning of LMW for sequencing
pJG008	TGTTATAGTTCCAATATTTTTGCCAAT	Forward primer for detection of mutation in Traes_7BL_F621D9B9E (MADS box protein)

pJG009	TTGAGCGATGCACAAAGC	Reverse primer for detection of mutation in Traes_7BL_F621D9B9E (MADS box protein)
pJG010	CGCCTCGTATTTGATGTTCA	Forward primer for detection of mutation in Traes_2AS_D659E88E9.1 (uncharacterized gene)
pJG011	GTAACAGCTTGCCGATGGAC	Reverse primer for detection of mutation in Traes_2AS_D659E88E9.1 (uncharacterized gene)
pJG012	GCCTCCTTCCTGATGTTC	Forward primer for detection of mutation in Traes_2AS_8FCC59363.1 (uncharacterized gene)
pJG013	TGCCGGTGTACACTTCTAGT	Reverse primer for detection of mutation in Traes_2AS_8FCC59363.1 (uncharacterized gene)
pJG014	ATGAAGACCTTAYTCATCC	Forward primer for cloning of gamma-gliadins for sequencing
pJG015	ACATACACGTTGCACATG	Reverse primer for cloning of gamma-gliadins for sequencing
pJG020	TTTGTCCTCCTTGCCATGGC	Forward primer for cloning of omega-gliadins for sequencing
pJG021	ATACTTATAACGTCGCTCCCAGAT	Reverse primer for cloning of omega-gliadins for sequencing
pJG022	TCATTGGCCACCGATGCTT	Reverse primer for cloning of omega-gliadins for sequencing

Supplementary Table 2. Illumina sequencing of alpha-gliadins in 18 T1 bread and durum wheat transgenic lines

				No. of reads				Frequency (%)			
Genotype	**Construct**	**T0 plant**	**T1 plant**	**Filtered reads**	**Insertions**	**Deletions**	**Total indels**	**Insertions**	**Deletions**	**Total indels**	**Cas9**
BW208 wt	**NA**	**NA**	**NA**	**211.415**	**0**	**0**	**0**	**0**	**0**	**0**	**NA**
	sgAlpha-2	10	T544	153.087	29.156	66.216	95.372	19,05	43,25	62,30	+
	sgAlpha-2	10	T545	163.004	20.564	101.903	122.467	12,62	62,52	75,13	+
	sgAlpha-2	10	T553	178.498	27.595	97.306	124.901	15,46	54,51	69,97	+
	sgAlpha-1	14	T567	105.983	0	1.058	1.058	0	1,00	1,00	+
	sgAlpha-1	14	T573	123.020	0	3.981	3.981	0	3,24	3,24	+
THA53 wt	**NA**	**NA**	**NA**	**196.316**	**0**	**0**	**0**	**0**	**0**	**0**	**NA**
	sgAlpha-2	20	V464	254.732	0	15.179	15.179	0	5,96	5,96	-
	sgAlpha-2	20	V465	126.518	112	9.835	9.947	0,09	7,77	7,86	+
	sgAlpha-2	20	V467	141.935	0	7.346	7.346	0	5,18	5,18	+
	sgAlpha-2	20	V468	279.894	0	14.450	14.450	0	5,16	5,16	-
DP wt	**NA**	**NA**	**NA**	**258.339**	**0**	**0**	**0**	**0**	**0**	**0**	**NA**
	sgAlpha-2	2	V525	192.189	457	22.040	22.497	0,24	11,47	11,71	-
	sgAlpha-2	2	V528	262.623	972	21.401	22.373	0,37	8,15	8,52	-
	sgAlpha-2	2	T666	144.275	0	15.722	15.722	0	10,90	10,90	+
	sgAlpha-2	2	T670	115.822	1.276	5.660	6.936	1,10	4,89	5,99	-
	sgAlpha-2	5	V520	264.423	0	39.047	39.047	0	14,77	14,77	-
	sgAlpha-2	5	T654	138.007	19.027	305	19.332	13,79	0,22	14,01	+
	sgAlpha-2	32	V511	152.378	0	2.414	2.414	0	1,58	1,58	+
	sgAlpha-2	32	V517	154.244	0	4.461	4.461	0	2,89	2,89	+

Supplementary Table 3. Gliadin and glutenin contents, total prolamin content, and gliadin to glutenin ratio of transgenic and wild-type T1 half-seeds from T0 lines.

Genotype	**Construct**	**Line**	**Plant Id**	ω-gliadins	**α**-gliadins	ɣ-gliadins	Total gliadin	HMW	LMW	Total glutenin	Prolamin	Gli/Glu
BW208	**NA**	**BW208**	**wt**	14,6	37,4	36,6	88,6	10,3	23,1	33,4	122,0	2,734
	sgAlpha-1	B686-6-C1-1	44	14,7	**12.2***	**17.4***	**44.3***	**25.7***	26,9	**52.7***	**96.9***	**0.850***
		B686-10-C3-1	14	**21.6***	**24.1***	**27.8***	73,6	**21.6***	25,5	**47.2***	120,7	**1.819***
		B686-10-C1-1	28	**39.3***	27,8	33,1	100,3	**26.6***	25,2	**51.8***	**152.1***	2,377
		B686-6-C1-2	52B	**36.2***	32,4	37,7	106,4	**24.5***	21,8	46,2	152,6	2,323
	sgAlpha-2	B683-3-C2-1	12	23,1	**12.7***	**18.6***	54,4	**31.3***	23,3	54,6	109,1	**0.988***
		B683-3-C4-1	10	9,2	**4.9***	**2.1***	**16.1***	**41.9***	**13.2***	**55.1***	**71.3***	**0.313***
		B683-3-C6-1	48	17,3	**13.6***	**21.1***	**52.0***	**24.8***	31,5	**56.4***	108,3	**0.925***
		B683-3-C5-2	49	16,9	**14.7***	**24.1***	**55.7***	**24.2***	**31.8***	**56.0***	111,7	**1.025***
THA53	**NA**	**THA53**	**wt**	13,2	38,4	35,3	86,9	12,4	23,5	35,9	122,8	2,598
	sgAlpha-1	B689-11-C1-1	77	**35.0***	**24.3***	42,2	101,5	**26.7***	18,6	45,3	146,8	2,241
	sgAlpha-2	B677-6-C1-1	20	14,4	**16.0***	**19.2***	**49.6***	**24.0***	18,4	42,4	**92.0***	**1.199***
		B677-6-C3-1	6	14,9	**7.8***	**10.9***	**33.6***	**30.7***	30,8	**61.5***	95,0	**0.569***
		B680-1-C1-1	15	15,7	**20.2***	**21.0***	**56.9***	**30.1***	31,9	**62.1***	118,9	**1.058***
		B680-1-C1-3	17	11,9	**9.9***	**16.3***	**38.0***	**25.2***	30,6	55,7	93,8	**0.681***
		B680-1-C2-1	21	15,6	**14.0***	**21.2***	**50.9***	**34.0***	22,0	**55.9***	106,8	**0.928***
DP	**NA**	**DP**	**wt**	8,3	35,2	33,0	76,5	5,6	27,0	32,6	109,1	2,423
	sgAlpha-2	B680-5-C1-1	32	6,9	**9.8***	**7.7***	**24.4***	**9.5***	**17.8***	27,3	**51.7***	**0.899***
		B681-2-C1-1	5	8,4	**10.4***	**13.0***	**31.9***	**11.1***	20,6	31,7	**63.7***	**1.200***
		B681-2-C1-2	2	6,3	**8.4***	**9.4***	**24.2***	**9.6***	24,4	34,0	**58.2***	**0.729***
		B681-5-C1-1	53	7,8	**15.2***	**10.7***	**33.7***	9,6	**39.7***	**49.3***	**82.9***	**0.695***

Gliadin and glutenin fractions were determined by RP-HPLC and expressed as µg/mg flour. Values for protein fraction are the mean of 10 grains from each T0 line. H MW, high molecular weight; LMW, low molecular weight; NA, Non applicable; wt, wild type.

* Means are significantly different to wild types as determined by Dunnett's multiple comparisons at $P < 0.05$.

Supplementary Table 4. Illumina sequencing of alpha-gliadins in 29 T2 bread and durum wheat transgenic lines

					No. of reads				Frequency (%)			
Genotype	**Construct**	**T0 plant**	**T1 plant**	**T2 plant**	**Filtered reads**	**Insertions**	**Deletions**	**Total indels**	**Insertions**	**Deletions**	**Total indels**	**Cas9**
BW208 wt	**NA**	**NA**	**NA**	**NA**	**244.228**	**0**	**0**	**0**	**0,00**	**0,00**	**0,00**	**NA**
	sgAlpha-2	10	T544	V601	247.674	18.350	145.801	164.151	7,41	58,87	66,28	-
	sgAlpha-2	10	T544	V603	213.965	19.471	123.189	142.660	9,10	57,57	66,67	-
	sgAlpha-2	10	T545	V723	173.544	8.817	106.803	115.620	5,08	61,54	66,62	+
	sgAlpha-2	10	T545	V726	218.717	11.200	131.426	142.626	5,12	60,09	65,21	+
	sgAlpha-2	10	T553	V657	232.478	5.494	140.636	146.130	2,36	60,49	62,86	-
	sgAlpha-2	10	T553	V660	281.734	5.516	177.520	183.036	1,96	63,01	64,97	-
	sgAlpha-2	10	T549	V581	252.645	5.089	137.775	142.864	2,01	54,53	56,55	-
	sgAlpha-2	12	T557	V701	292.782	0	2.545	2.545	0,00	0,87	0,87	+
	sgAlpha-2	12	T557	V704	172.220	27	7.023	7.050	0,02	4,08	4,09	-
	sgAlpha-2	12	T557	V705	231.067	0	1.761	1.761	0,00	0,76	0,76	-
	sgAlpha-2	12	T559	V733	223.210	0	902	902	0,00	0,40	0,40	+
	sgAlpha-2	12	T559	V738	219.081	0	997	997	0,00	0,46	0,46	+
	sgAlpha-2	12	T559	V739	197.923	0	824	824	0,00	0,42	0,42	-
	sgAlpha-2	12	T559	V740	218.499	0	851	851	0,00	0,39	0,39	-
	sgAlpha-1	14	T567	V631	122.997	0	2.892	2.896	0,00	2,35	2,35	+
	sgAlpha-1	14	T567	V634	118.898	246	4.427	4.673	0,21	3,72	3,93	+
	sgAlpha-1	14	T573	V641	157.236	61	1.958	2.019	0,04	1,25	1,28	+
	sgAlpha-1	14	T573	V644	181.483	128	1.808	1.936	0,07	1,00	1,07	+
THA53 wt	**NA**	**NA**	**NA**	**NA**	**196.316**	**0**	**0**	**0**	**0,00**	**0,00**	**0,00**	**NA**
	sgAlpha-2	6	T450	V791	215.380	0	2.841	2.841	0,00	1,32	1,32	+
DP wt	**NA**	**NA**	**NA**	**NA**	**380.497**	**0**	**0**	**0**	**0,00**	**0,00**	**0,00**	**NA**
	sgAlpha-2	2	T666	V773	124.504	0	29.778	29.778	0,00	23,92	23,92	+
	sgAlpha-2	2	T666	V775	248.952	0	43.119	43.119	0,00	17,32	17,32	-
	sgAlpha-2	2	T666	V778	157.464	0	37.633	37.633	0,00	23,90	23,90	-
	sgAlpha-2	2	T666	V780	270.513	0	27.929	27.929	0,00	10,32	10,32	-
	sgAlpha-2	2	T670	V752	152.753	1.591	7.000	8.591	1,04	4,58	5,62	-
	sgAlpha-2	2	T670	V756	232.321	2.247	11.119	13.366	0,97	4,79	5,75	-
	sgAlpha-2	2	T670	V759	182.702	2.243	9.106	11.349	1,23	4,98	6,21	-
	sgAlpha-2	5	T654	V768	183.188	16.688	4.462	21.150	9,11	2,44	11,55	+
	sgAlpha-2	5	V520	X458	167.940	0	20.201	20.201	0,00	12,03	12,03	-
	sgAlpha-2	5	V520	X459	196.867	0	22.389	22.389	0,00	11,37	11,37	-

Supplementary Table 5. Gliadin and glutenin contents, total prolamin content, gliadin to glutenin ratio, and SDS sedimentation test of transgenic and wild-type T2 seeds from T1 lines

Genotyp	Construct	T0 plant id	T1 plant	ω-gliadins	α-gliadins	γ-gliadins	Total gliadir	HMW	LMW	Total gluten	Prolamins	Gli/Glu
BW208	**NA**	**wt**	**wt**	14,3	35,2	37,7	87,2	10,1	22,6	32,6	119,9	2,702
	sgAlpha-1	14	T566	**30.9***	**14.5***	**26.4***	**71.9***	22,2	22,0	44,1	116,0	**1.633***
			T567	**31.4***	**13.6***	**26.6***	**71.6***	**34.4***	27,9	**62.2***	133,8	**1.276***
			T571	**29.8***	**17.5***	**27.0***	**74.3***	**38.2***	27,9	**66.1***	**140.4***	**1.156***
			T572	**35.1***	**16.5***	**32.5***	84,1	**44.4***	28,9	**73.3***	**157.4***	**1.154***
			T573	**37.7***	**13.8***	**26.1***	77,6	**33.6***	25,2	**58.8***	136,4	**1.327***
	sgAlpha-2	10	T544	13,6	**4.1***	**6.3***	**24.1***	19,6	**9.1***	28,6	**52.7***	**0.859***
			T545	13,5	**4.5***	**4.2***	**22.2***	**21.8***	**6.8***	28,6	**50.8***	**0.786***
			T546	14,0	**3.6***	**4.4***	**22.0***	**24.4***	**10.7***	35,2	**57.2***	**0.641***
			T547	15,4	**4.7***	**6.0***	**26.1***	**46.3***	**11.7***	**57.9***	**84.0***	**0.451***
			T548	11,5	**3.4***	**3.6***	**18.5***	**32.1***	**6.5***	38,6	**57.1***	**0.484***
			T549	16,5	**5.8***	**6.5***	**28.8***	**37.2***	**12.2***	**49.3***	**78.2***	**0.647***
			T550	13,0	**2.4***	**4.3***	**19.6***	**23.7***	**9.5***	33,2	**52.8***	**0.599***
			T551	16,3	**6.6***	**6.4***	**29.2***	**53.9***	**13.5***	**67.4***	**96.6***	**0.445***
			T552	13,0	**4.4***	**3.9***	**21.2***	**38.6***	**9.3***	47,9	**69.1***	**0.473***
			T553	17,9	**5.9***	**5.6***	**29.4***	**23.3***	**7.6***	30,9	**60.2***	**0.956***
		12	T558	17,0	**8.1***	**13.6***	**38.6***	**31.5***	20,5	**52.0***	**90.7***	**0.849***
			T559	17,1	**12.3***	**21.9***	**51.9***	**31.2***	25,3	**56.5***	108,4	**0.931***
			T561	**20.6***	**8.6***	**15.0***	**44.3***	**25.3***	10,8	36,0	**80.3***	**1.305***
THA53	**NA**	**wt**	**wt**	14,5	34,1	34,9	83,5	14,0	25,8	39,8	123,3	2,124
	sgAlpha-2	6	T450	**18.3***	**11.2***	**24.1***	**53.5***	18,6	**18.1***	36,7	**90.1***	**1.470***
			T454	16,2	**13.1***	**25.2***	**54.5***	**42.6***	27,9	**70.5***	125,0	**0.795***
			T455	**19.5***	**10.9***	**24.9***	**55.3***	**22.9***	**18.8***	41,7	**97.0***	**1.337***
DP	**NA**	**wt**	**wt**	11,0	31,5	33,6	76,0	6,4	29,7	36,1	112,0	2,170
	Alpha-2	5	T654	**18.9***	**20.6***	36,5	76,0	**21.5***	40,2	**61.7***	121,9	**1.417***
		2	T666	**17.8***	**11.2***	30,5	**59.5***	**30.1***	31,8	**61.8***	109,3	**0.799***
			T668	13,1	**5.9***	**23.8***	**42.9***	**30.5***	28,7	**59.2***	108,2	**0.850***
			T670	**25.1***	**8.6***	40,1	73,8	**17.5***	34,2	**51.7***	107,8	**1.110***

Gliadin and glutenin fractions were determined by RP-HPLC and expressed as µg/mg flour. Values for protein fraction are the mean of 5-10 hal-seeds from each T1 line. HMW, high molecular weight; LMW, low molecular weight; NA, Non applicable; ND, data not determined; wt, wild type.

* Means are significantly different to wild types as determined by Dunnett's multiple comparisons at $P < 0.05$.

Supplementary Table 6. Gliadin and glutenin contents, total prolamin content, gliadin to glutenin ratio, and SDS sedimentation test of transgenic and wild-type T3 seeds from T2 lines

Genotype	Construct	T0 plant id	T1 plant	ω-gliadins	α-gliadins	γ-gliadins	otal gliadins	HMW	LMW	otal glutenins	Prolamins	Gli/Glu
BW208	**NA**	**wt**	**wt**	12,9	38,0	30,4	81,3	15,6	29,3	44,9	126,2	1,809
	sgAlpha-1	14	T567	**18.3***	**15.0***	**24.4***	**57.7***	**28.9***	29,7	58,5	116,2	**0.985***
			T573	**19.0***	**14.9***	**23.3***	**57.2***	**35.9***	34,6	**70.5***	127,8	**0.810***
	sgAlpha-2	10	T544	**8.8***	**6.1***	**4.6***	**19.6***	20,9	**13.5***	34,4	54,0	**0.568***
			T545	9,9	**5.5***	**3.6***	**19.0***	**26.2***	**11.3***	37,5	56,6	**0.507***
			T546	**7.2***	**3.9***	**2.6***	**13.5***	20,3	**7.7***	28,0	41,6	**0.482***
			T549	11,2	**6.8***	**6.4***	**24.3***	**34.1***	**13.2***	47,4	71,7	**0.513***
			T550	10,7	**6.5***	**4.9***	**22.1***	**35.7***	**12.6***	48,4	70,4	**0.456***
			T553	12,6	**6.9***	**5.7***	**25.3***	**38.9***	**13.2***	52,1	77,4	**0.485***
		12	T558	15,9	**10.4***	**11.9***	**38.3***	**41.2***	24,4	**65.6***	103,9	**0.583***
			T559	17,8	**27.0***	**23.4***	**68.2***	**32.4***	34,5	**66.8***	135,0	**1.019***
			T561	15,6	**10.1***	**11.2***	**36.9***	**33.1***	17,7	50,9	87,8	**0.725***
THA53	**NA**	**wt**	**wt**	11,8	42,9	32,3	86,9	17,7	37,0	54,7	141,6	1,649
	sgAlpha-2	6	T450	13,9	**16.7***	**20.8***	**51.5***	**28.7***	19,0	47,7	**99.2***	1,082
			T455	12,6	**11.9***	**17.8***	**42.3***	**29.8***	23,6	53,3	**95.6***	**0.793***
DP	**NA**	**wt**	**wt**	7,4	35,5	26,8	69,6	8,9	42,9	51,8	121,4	1,370
	Alpha-2	5	T654	**11.5***	**18.2***	25,8	**55.5***	**20.3***	30,7	51,0	**106.5***	1,089
		2	T666	7,9	**6.4***	**14.1***	**28.4***	**23.4***	**25.7***	49,1	**77.5***	**0.579***
			T670	**12.0***	**8.1***	28,4	**48.5***	**17.7***	39,9	57,6	**106.0***	**0.842***

Gliadin and glutenin fractions were determined by RP-HPLC and expressed as µg/mg flour. T2 lines from Supplementary Table 5 were multiplied and equivalent amounts of grains from each line were bulked and milled for protein determination. Values for each protein fraction are the mean of four replications. HMW, high molecular weight; LMW, low molecular weight; NA, Non applicable; ND, data not determined; wt, wild type.

* Means are significantly different to wild types as determined by Dunnett's multiple comparisons at $P < 0.05$.

References

Barro, F., Iehisa, J.C.M., Giménez, M.J., García-Molina, M.D., Ozuna, C.V., Comino, I., et al. (2016). Targeting of prolamins by RNAi in bread wheat: effectiveness of seven silencing-fragment combinations for obtaining lines devoid of coeliac disease epitopes from highly immunogenic gliadins. *Plant Biotechnology Journal* 14(3), 986-996. doi: 10.1111/pbi.12455.

Edgar, R.C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26(19), 2460-2461. doi: 10.1093/bioinformatics/btq461.

Edgar, R.C., and Flyvbjerg, H. (2015). Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics* 31(21), 3476-3482. doi: 10.1093/bioinformatics/btv401.

Gil-Humanes, J., Pistón, F., Altamirano-Fortoul, R., Real, A., Comino, I., Sousa, C., et al. (2014a). Reduced-Gliadin Wheat Bread: An Alternative to the Gluten-Free Diet for Consumers Suffering Gluten-Related Pathologies. *PLOS ONE* 9(3), e90898. doi: 10.1371/journal.pone.0090898.

Gil-Humanes, J., Pistón, F., Barro, F., and Rosell, C.M. (2014b). The Shutdown of Celiac Disease-Related Gliadin Epitopes in Bread Wheat by RNAi Provides Flours with Increased Stability and Better Tolerance to Over-Mixing. *PLOS ONE* 9(3), e91931. doi: 10.1371/journal.pone.0091931.

Gil-Humanes, J., Pistón, F., Rosell, C.M., and Barro, F. (2012). Significant down-regulation of γ-gliadins has minor effect on gluten and starch properties of bread wheat. *Journal of Cereal Science* 56(2), 161-170. doi: 10.1016/j.jcs.2012.02.009.

Gil-Humanes, J., Pistón, F., Shewry, P.R., Tosi, P., and Barro, F. (2011). Suppression of gliadins results in altered protein body morphology in wheat. *Journal of Experimental Botany* 62(12), 4203-4213. doi: 10.1093/jxb/err119.

Gil-Humanes, J., Pistón, F., Tollefsen, S., Sollid, L.M., and Barro, F. (2010). Effective shutdown in the expression of celiac disease-related wheat gliadin T-cell epitopes by RNA interference. *PNAS* 107(39), 7023–17028. doi: 10.1073/pnas.1007773107.

Katoh, K., Misawa, K., Kuma, K.i., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* 30(14), 3059-3066. doi: 10.1093/nar/gkf436.

Mann, D.G.J., LaFayette, P.R., Abercrombie, L.L., King, Z.R., Mazarei, M., Halter, M.C., et al. (2012). Gateway-compatible vectors for high-throughput gene functional analysis in switchgrass (Panicum virgatum L.) and other monocot species. *Plant Biotechnology Journal* 10(2), 226-236. doi: 10.1111/j.1467-7652.2011.00658.x.

Mustalahti, K., Catassi, C., Reunanen, A., Fabiani, E., Heier, M., McMillan, S., et al. (2010). The prevalence of celiac disease in Europe: Results of a centralized, international mass screening project. *Annals of Medicine* 42(8), 587-595. doi: 10.3109/07853890.2010.505931.

Ozuna, C.V., Iehisa, J.C.M., Giménez, M.J., Alvarez, J.B., Sousa, C., and Barro, F. (2015). Diversification of the celiac disease α-gliadin complex in wheat: a 33-mer peptide with six overlapping epitopes, evolved following polyploidization. *Plant Journal* 82(5), 794-805. doi: 10.1111/tpj.12851.

Pistón, F., Gil-Humanes, J., Rodríguez-Quijano, M., and Barro, F. (2011). Down-regulating γ-gliadins in bread wheat leads to non-specific increases in other gluten proteins and has no major effect on dough gluten strength. *PloS one* 6(9), e24754-e24754. doi: 10.1371/journal.pone.0024754.

Pistón, F., Marín, S., Hernando, A., and Barro, F. (2009). Analysis of the activity of a γ-gliadin promoter in transgenic wheat and characterization of gliadin synthesis in wheat by MALDI-TOF during grain development. *Molecular breeding* 23(4), 655-667. doi: 10.1007/s11032-009-9263-1.

Price, M.N., Dehal, P.S., and Arkin, A.P. (2010). FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLOS ONE* 5(3), e9490. doi: 10.1371/journal.pone.0009490.

Sapone, A., Lammers, K.M., Casolaro, V., Cammarota, M., Giuliano, M.T., De Rosa, M., et al. (2011). Divergence of gut permeability and mucosal immune gene expression in two gluten-associated conditions: celiac disease and gluten sensitivity. *BMC Medicine* 9(1), 23. doi: 10.1186/1741-7015-9-23.

Tye-Din, J.A., Stewart, J.A., Dromey, J.A., Beissbarth, T., van Heel, D.A., Tatham, A., et al. (2010). Comprehensive, Quantitative Mapping of T Cell Epitopes in Gluten in Celiac Disease. *Science Translational Medicine* 2(41), 41ra51-41ra51. doi: 10.1126/scitranslmed.3001012.

Valdés, I., García, E., Llorente, M., and Méndez, E. (2003). Innovative approach to low-level gluten determination in foods using a novel sandwich enzyme-linked immunosorbent assay protocol. *European Journal of Gastroenterology & Hepatology* 15(5), 465-747. doi: 10.1097/01.meg.0000059119.41030.df.

Williams, P. (1988). Crop Quality Evaluation Methods and Guidelines*. International Center for Agricultural Research in the Dry Areas.*

CHAPTER 5

Conclusions

CONCLUSIONS

i) Tritordeum may be tolerated by at least a sub-set of NCGS sufferers who do not require strict exclusion of gluten from their diet, without suffering negative effects on gastro-intestinal health (Chapter 2).

ii) Tritordeum consumption promotes a trend to increase the relative abundance of some SCFA-producing bacteria, especially the butyrate-producing genus, *Faecalibacterium*. This may suggest an increase of the intestinal healthy status in these patients, as butyrate contribute to maintenance of the intestinal barrier, promoting the intestinal homeostasis and have anti-inflammatory effects (Chapter 2).

iii) PBMCs proliferation assay and INF-γ release provided three wheat RNAi lines whose stimulatory response does not differ from rice protein extract uses as a gluten free negative control. Optimal protein composition is a pronounced decrease in the α-gliadins, containing very low or no DQ2.5 and p31-43 α-gliadin epitopes, and low G12 values. (Chapter 3).

iv) The downregulation by RNAi of prolamin fractions is an excellent tool to develop wheat varieties with very low content of CD immunogenic epitopes, particularly α-gliadins, the most immunogenic fraction known so far (Chapter 3).

v) PBMCs provide a useful tool for the validation of stimulatory capacity of RNAi wheat lines differing in the protein composition triggering immune response in CD (Chapter 3).

vi) High mutation frequency and specificity can be achieved using CRISPR/Cas9 to modify complex genomic loci such as the α-gliadin gene family in bread and durum wheat (Chapter 4).

vii) Reduction in the content of α-gliadins by CRISPR-Cas9 technology in the seed kernel provided bread and durum wheat lines with reduced immunoreactivity for gluten-intolerant consumers (Chapter 4).

viii) The low-gluten, transgene-free wheat lines constitute an unprecedented advance, and could be used to produce low gluten foodstuff and serve as source material in plant breeding programs to introgress this trait into elite wheat varieties (Chapter 4).